

## PATENT OPERATION TREATY

PCT

## NOTIFICATION OF ELECTION

(PCT Rule 61.2)

From the INTERNATIONAL BUREAU

To:

United States Patent and Trademark  
Office  
(Box PCT)  
Crystal Plaza 2  
Washington, DC 20231  
ÉTATS-UNIS D'AMÉRIQUE

in its capacity as elected Office

Date of mailing (day/month/year)

17 May 1999 (17.05.99)

International application No.

PCT/AU98/00743

Applicant's or agent's file reference

International filing date (day/month/year)

11 September 1998 (11.09.98)

Priority date (day/month/year)

12 September 1997 (12.09.97)

Applicant

LI, Zhongyi et al

1. The designated Office is hereby notified of its election made:

☒ in the demand filed with the International Preliminary Examining Authority on:

06 April 1999 (06.04.99)

☐ in a notice effecting later election filed with the International Bureau on:2. The election ☒ was☐ was not

made before the expiration of 19 months from the priority date or, where Rule 32 applies, within the time limit under Rule 32.2(b).

The International Bureau of WIPO  
34, chemin des Colombettes  
1211 Geneva 20, Switzerland

Facsimile No.: (41-22) 740.14.35

Authorized officer

S. Mafla

Telephone No.: (41-22) 338.83.38

## PATENT COOPERATION TREATY

PCT

NOTIFICATION OF THE RECORDING  
OF A CHANGE(PCT Rule 92bis.1 and  
Administrative Instructions, Section 422)

From the INTERNATIONAL BUREAU

To:

GRIFFITH HACK  
509 St. Kilda Road  
Melbourne, VIC 3004  
AUSTRALIE

Date of mailing (day/month/year)

02 November 1999 (02.11.99)

Applicant's or agent's file reference

## IMPORTANT NOTIFICATION

International application No.

PCT/AU98/00743

International filing date (day/month/year)

11 September 1998 (11.09.98)

## 1. The following indications appeared on record concerning:



the applicant



the inventor



the agent



the common representative

Name and Address

THE AUSTRALIAN NATIONAL UNIVERSITY  
Acton, ACT 2601  
AU

State of Nationality

AU

State of Residence

AU

Telephone No.

Facsimile No.

Teleprinter No.

## 2. The International Bureau hereby notifies the applicant that the following change has been recorded concerning:



the person



the name



the address



the nationality



the residence

Name and Address

State of Nationality

State of Residence

Telephone No.

Facsimile No.

Teleprinter No.

## 3. Further observations, if necessary:

**The applicant above-identified should now be deleted from record.**

## 4. A copy of this notification has been sent to:



the receiving Office



the International Searching Authority



the International Preliminary Examining Authority



the designated Offices concerned



the elected Offices concerned



other:

The International Bureau of WIPO  
34, chemin des Colombettes  
1211 Geneva 20, Switzerland

Authorized officer

Eugénia Santos

Facsimile No.: (41-22) 740.14.35

Telephone No.: (41-22) 338.83.38

**PATENT COOPERATION TREATY**  
**PCT**  
**INTERNATIONAL PRELIMINARY EXAMINATION REPORT**  
(PCT Article 36 and Rule 70)

Applicant's or agent's file reference VS:SJB:LM:FP10105	<b>FOR FURTHER ACTION</b>	See Notification of Transmittal of International Preliminary Examination Report (Form PCT/IPEA/416).
International application No.  PCT/AU 98/00743	International filing date (day/month/year)  11 September 1998	Priority Date (day/month/year)  12 September 1997
International Patent Classification (IPC) or national classification and IPC  Int. Cl. <sup>6</sup> C12N 9/24, 15/55		
Applicant  COMMONWEALTH SCIENTIFIC AND INDUSTRIAL RESEARCH ORGANISATION		

1. This international preliminary examination report has been prepared by this International Preliminary Examining Authority and is transmitted to the applicant according to Article 36.
2. This REPORT consists of a total of 3 sheets, including this cover sheet.
- ☒ This report is also accompanied by ANNEXES, i.e., sheets of the description, claims and/or drawings which have been amended and are the basis for this report and/or sheets containing rectifications made before this Authority (see Rule 70.16 and Section 607 of the Administrative Instructions under the PCT).
- These annexes consist of a total of 4 sheet(s).
3. This report contains indications relating to the following items:
- |      |                                     |   |
|------|-------------------------------------|---|
| I    | <input checked="" type="checkbox"/> | Basis of the report   |
| II   | <input type="checkbox"/>            | Priority  |
| III  | <input type="checkbox"/>            | Non-establishment of opinion with regard to novelty, inventive step and industrial applicability  |
| IV   | <input type="checkbox"/>            | Lack of unity of invention  |
| V    | <input checked="" type="checkbox"/> | Reasoned statement under Article 35(2) with regard to novelty, inventive step or industrial applicability; citations and explanations supporting such statement |
| VI   | <input type="checkbox"/>            | Certain documents cited   |
| VII  | <input type="checkbox"/>            | Certain defects in the international application  |
| VIII | <input type="checkbox"/>            | Certain observations on the international application   |

Date of submission of the demand 6 April 1999	Date of completion of the report 30 June 1999
Name and mailing address of the IPEA/AU AUSTRALIAN PATENT OFFICE PO BOX 200 WODEN ACT 2606 AUSTRALIA Facsimile No. (02) 6285 3929	Authorized Officer  <b>PHILIPPA WYRDEMAN</b>  Telephone No. (02) 6283 2554

**I Basis of the report****1. With regard to the elements of the international application:\***

- ☐ the international application as originally filed.
- ☒ the description, pages 1-6, 8-60 as originally filed,  
pages , filed with the demand, <sup>21</sup>  
pages 7 and 7a, filed with the letter of <sup>21</sup> 18 June 1999.
- ☒ the claims, pages 117, 199-122, as originally filed,  
pages , as amended (together with any statement) under Article 19,  
pages , filed with the demand, <sup>21</sup>  
pages 116 and 118, filed with the letter of <sup>21</sup> 18 June 1999.
- ☒ the drawings, pages 1/44-44/44 , as originally filed,  
pages , filed with the demand,  
pages , filed with the letter of .
- ☒ the sequence listing part of the description:  
pages 61-115 , as originally filed  
pages , filed with the demand  
pages , filed with the letter of

**2. With regard to the language, all the elements marked above were available or furnished to this Authority in the language in which the international application was filed, unless otherwise indicated under this item.**

These elements were available or furnished to this Authority in the following language which is:

- ☐ the language of a translation furnished for the purposes of international search (under Rule 23.1(b)).
- ☐ the language of publication of the international application (under Rule 48.3(b)).
- ☐ the language of the translation furnished for the purposes of international preliminary examination (under Rules 55.2 and/or 55.3).

**3. With regard to any nucleotide and/or amino acid sequence disclosed in the international application, was on the basis of the sequence listing:**

- ☒ contained in the international application in written form.
- ☐ filed together with the international application in computer readable form.
- ☐ furnished subsequently to this Authority in written form.
- ☐ furnished subsequently to this Authority in computer readable form.
- ☐ The statement that the subsequently furnished written sequence listing does not go beyond the disclosure in the international application as filed has been furnished.
- ☐ The statement that the information recorded in computer readable form is identical to the written sequence listing has been furnished

**4. ☐ The amendments have resulted in the cancellation of:**

- ☐ the description, pages
- ☐ the claims, Nos.
- ☐ the drawings, sheets/fig

**5. ☐ This report has been established as if (some of) the amendments had not been made, since they have been considered to go beyond the disclosure as filed, as indicated in the Supplemental Box (Rule 70.2(c)).\*\***

\* Replacement sheets which have been furnished to the receiving Office in response to an invitation under Article 14 are referred to in this report as "originally filed" and are not annexed to this report since they do not contain amendments (Rules 70.16 and 70.17).

\*\* Any replacement sheet containing such amendments must be referred to under item 1 and annexed to this report

**V. Reasoned statement under Article 35(2) with regard to novelty, inventive step or industrial applicability; citations and explanations supporting such statement****1. Statement**

Novelty (N)	Claims 1-52	YES
	Claims none	NO
Inventive step (IS)	Claims 1-52	YES
	Claims none	NO
Industrial applicability (IA)	Claims 1-52	YES
	Claims none	NO

**2. Citations and explanations (Rule 70.7)**

The closest prior art is D6 (Nair et al) as listed on the International Search Report. D6 discloses an N-terminal sequence specifically excluded from the claimed enzymes of the present application. The claims are thus considered both novel and inventive in light of the prior art.

The claimed matter is considered industrially applicable.

determined can also be used as probes for the identification and isolation of corresponding sequences, including promoter sequences, from other cereal plant species.

In its most general aspect, the invention provides  
5 a nucleic acid sequence encoding an enzyme of the starch biosynthetic pathway in a cereal plant, said enzyme being selected from the group consisting of starch branching enzyme I, starch branching enzyme II, starch soluble  
10 synthase I, and debranching enzyme, with the proviso that the enzyme is not soluble starch synthase I of rice, or starch branching enzyme I of rice or maize, and that starch branching enzyme II does not have the N-terminal amino acid sequence:

15 AASPGKVLVPDGEDDLASPA.

Preferably the nucleic acid sequence is a DNA sequence, and may be genomic DNA or cDNA. Preferably the sequence is one which is functional in wheat. More  
20 preferably the sequence is derived from a *Triticum* species, most preferably *Triticum tauschii*.

Where the sequence encodes soluble starch synthase, preferably the sequence encodes the 75 kD soluble starch synthase of wheat.

25 Biologically-active untranslated control sequences of genomic DNA are also within the scope of the invention. Thus the invention also provides the promoter of an enzyme as defined above.

In a preferred embodiment of this aspect of the  
30 invention, there is provided a nucleic acid construct comprising a nucleic acid sequence of the invention, a biologically-active fragment thereof, or a fragment thereof encoding a biologically-active fragment of an enzyme as defined above, operably linked to one or more nucleic acid  
35 sequences facilitating expression of said enzyme in a plant, preferably a cereal plant. The construct may be a plasmid or a vector, preferably one suitable for use in the transformation of a plant. A particularly suitable vector is a bacterium of the genus *Agrobacterium*, preferably

- 7a -

: *Agrobacterium tumefaciens*. Methods of transforming cereal plants using *Agrobacterium tumefaciens* are known; see for example Australian Patent No. 667939 by Japan Tobacco Inc.,

CLAIMS

1. A nucleic acid sequence encoding an enzyme of the starch biosynthetic pathway in a cereal plant, wherein the enzyme is selected from the group consisting of starch  
5 branching enzyme I, starch branching enzyme II, starch soluble synthase I, and debranching enzyme, with the proviso that the enzyme is not soluble starch synthase I of rice, or starch branching enzyme I of rice or maize, and that starch branching enzyme II does not have the N-terminal amino acid  
10 sequence:

AASPGKVLVPDGEDDLASPA.

2. A sequence according to claim 1, wherein the  
15 sequence is a genomic DNA or cDNA sequence.
3. A sequence according to claim 1 or claim 2, wherein the sequence is functional in wheat.
- 20 4. A sequence according to any one of claims 1 to 3, wherein the sequence is derived from a *Triticum* species.
5. A sequence according to claim 4, wherein the *Triticum* species is *Triticum tauschii*.
- 25 6. A sequence according to <sup>claim</sup>any one of claims 1 to 5, wherein the sequence encodes starch branching enzyme I or a biologically-active fragment thereof, and wherein the sequence has at least 70% sequence homology with the  
30 sequence shown in SEQ ID NO:5 or SEQ ID NO:9.
7. A sequence according to claim 6, wherein the homology is at least 90%.
- 35 8. A sequence according to any one of claims 1 to 5, wherein the sequence encodes starch branching enzyme II or a biologically-active fragment thereof, and wherein the sequence has at least 70% sequence homology with the sequence shown in SEQ ID NO:10.



- 118 -

biologically-active fragment thereof, and wherein the promoter sequence has at least 70% sequence homology with the sequence shown in SEQ ID No:8.

5 18. A sequence according to claim 17, wherein the homology is at least 90%.

19. A promoter according to claim 16, wherein the promoter is a starch soluble synthase I promoter or  
10 biologically-active fragment thereof, and wherein the promoter sequence has at least 70% sequence homology with the sequence shown in SEQ ID No:15.

20. A sequence according to claim 19, wherein the  
15 homology is at least 90%.

21. A nucleic acid construct comprising a nucleic acid sequence encoding an enzyme of the starch biosynthetic pathway in a cereal plant, operably linked to one or more  
20 nucleic acid sequences facilitating expression of the nucleic acid sequence in a plant, wherein the enzyme is selected from the group consisting of starch branching enzyme I, starch branching enzyme II, starch soluble synthase I, and debranching enzyme, with the proviso that  
25 the enzyme is not soluble starch synthase I of rice, or starch branching enzyme I of rice or maize, a biologically-active fragment thereof, and that starch branching enzyme II does not have the N-terminal amino acid sequence:

30 AASPGKVLVPDGEDDLASPA.

22. A nucleic acid construct for targeting a gene to the endosperm of a cereal plant, comprising one or more promoter sequences selected from the group consisting of  
35 SBE I promoter, SBE II promoter, SSS I promoter, and DBE promoter, operatively linked to a nucleic acid sequence encoding a protein, wherein the expression of the targetted gene in the endosperm of a cereal plant is modified.

REPLACED BY  
ART 34 AMDT

PATENT COOPERATION TREATY

PCT

INTERNATIONAL PRELIMINARY EXAMINATION REPORT

19  
REC'D 13 JUL 1999

REPORT PCT

(PCT Article 36 and Rule 70)

Applicant's or agent's file reference VS:SJB:LM:FP10105	<b>FOR FURTHER ACTION</b>	See Notification of Transmittal of International Preliminary Examination Report (Form PCT/IPEA/416).
International application No.  PCT/AU 98/00743	International filing date (day/month/year)  11 September 1998	Priority Date (day/month/year)  12 September 1997
International Patent Classification (IPC) or national classification and IPC  Int. Cl. <sup>6</sup> C12N 9/24, 15/55		
Applicant  COMMONWEALTH SCIENTIFIC AND INDUSTRIAL RESEARCH ORGANISATION		

1.	This international preliminary examination report has been prepared by this International Preliminary Examining Authority and is transmitted to the applicant according to Article 36.
2.	This REPORT consists of a total of 3 sheets, including this cover sheet.  <input checked="" type="checkbox"/> This report is also accompanied by ANNEXES, i.e., sheets of the description, claims and/or drawings which have been amended and are the basis for this report and/or sheets containing rectifications made before this Authority (see Rule 70.16 and Section 607 of the Administrative Instructions under the PCT).  These annexes consist of a total of 4 sheet(s).
3.	This report contains indications relating to the following items:  I <input checked="" type="checkbox"/> Basis of the report II <input type="checkbox"/> Priority III <input type="checkbox"/> Non-establishment of opinion with regard to novelty, inventive step and industrial applicability IV <input type="checkbox"/> Lack of unity of invention V <input checked="" type="checkbox"/> Reasoned statement under Article 35(2) with regard to novelty, inventive step or industrial applicability; citations and explanations supporting such statement VI <input type="checkbox"/> Certain documents cited VII <input type="checkbox"/> Certain defects in the international application VIII <input type="checkbox"/> Certain observations on the international application

Date of submission of the demand 6 April 1999	Date of completion of the report 30 June 1999
Name and mailing address of the IPEA/AU AUSTRALIAN PATENT OFFICE PO BOX 200 WODEN ACT 2606 AUSTRALIA Facsimile No. (02) 6285 3929	Authorized Officer  <b>PHILIPPA WYRDEMAN</b> Telephone No. (02) 6283 2554

**I. Basis of the report****1. With regard to the elements of the international application:\***

- ☐ the international application as originally filed.
- ☒ the description, pages 1-6, 8-60 as originally filed,  
pages , filed with the demand, <sup>21</sup>  
pages 7 and 7a, filed with the letter of <sup>21</sup> 18 June 1999.
- ☒ the claims, pages 117, 199-122, as originally filed,  
pages , as amended (together with any statement) under Article 19,  
pages , filed with the demand,  
pages 116 and 118, filed with the letter of <sup>21</sup> 18 June 1999.
- ☒ the drawings, pages 1/44-44/44 , as originally filed,  
pages , filed with the demand,  
pages , filed with the letter of .
- ☒ the sequence listing part of the description:  
pages 61-115 , as originally filed  
pages , filed with the demand  
pages , filed with the letter of

**2. With regard to the language, all the elements marked above were available or furnished to this Authority in the language in which the international application was filed, unless otherwise indicated under this item.**

These elements were available or furnished to this Authority in the following language which is:

- ☐ the language of a translation furnished for the purposes of international search (under Rule 23.1(b)).
- ☐ the language of publication of the international application (under Rule 48.3(b)).
- ☐ the language of the translation furnished for the purposes of international preliminary examination (under Rules 55.2 and/or 55.3).

**3. With regard to any nucleotide and/or amino acid sequence disclosed in the international application, was on the basis of the sequence listing:**

- ☒ contained in the international application in written form.
- ☐ filed together with the international application in computer readable form.
- ☐ furnished subsequently to this Authority in written form.
- ☐ furnished subsequently to this Authority in computer readable form.
- ☐ The statement that the subsequently furnished written sequence listing does not go beyond the disclosure in the international application as filed has been furnished.
- ☐ The statement that the information recorded in computer readable form is identical to the written sequence listing has been furnished

**4. ☐ The amendments have resulted in the cancellation of:**

- ☐ the description, pages
- ☐ the claims, Nos.
- ☐ the drawings, sheets/fig

**5. ☐ This report has been established as if (some of) the amendments had not been made, since they have been considered to go beyond the disclosure as filed, as indicated in the Supplemental Box (Rule 70.2(c)).\*\***

\* Replacement sheets which have been furnished to the receiving Office in response to an invitation under Article 14 are referred to in this report as "originally filed" and are not annexed to this report since they do not contain amendments (Rules 70.16 and 70.17).

\*\* Any replacement sheet containing such amendments must be referred to under item 1 and annexed to this report

**V. Reasoned statement under Article 35(2) with regard to novelty, inventive step or industrial applicability; citations and explanations supporting such statement****1. Statement**

Novelty (N)	Claims 1-52	YES
	Claims none	NO
Inventive step (IS)	Claims 1-52	YES
	Claims none	NO
Industrial applicability (IA)	Claims 1-52	YES
	Claims none	NO

**2. Citations and explanations (Rule 70.7)**

The closest prior art is D6 (Nair et al) as listed on the International Search Report. D6 discloses an N-terminal sequence specifically excluded from the claimed enzymes of the present application. The claims are thus considered both novel and inventive in light of the prior art.

The claimed matter is considered industrially applicable.

- 7 -

determined can also be used as probes for the identification and isolation of corresponding sequences, including promoter sequences, from other cereal plant species.

In its most general aspect, the invention provides  
5 a nucleic acid sequence encoding an enzyme of the starch biosynthetic pathway in a cereal plant, said enzyme being selected from the group consisting of starch branching enzyme I, starch branching enzyme II, starch soluble synthase I, and debranching enzyme, with the proviso that  
10 the enzyme is not soluble starch synthase I of rice, or starch branching enzyme I of rice or maize.

Preferably the nucleic acid sequence is a DNA sequence, and may be genomic DNA or cDNA. Preferably the sequence is one which is functional in wheat. More  
15 preferably the sequence is derived from a *Triticum* species, most preferably *Triticum tauschii*.

Where the sequence encodes soluble starch synthase, preferably the sequence encodes the 75 kD soluble starch synthase of wheat.

20 Biologically-active untranslated control sequences of genomic DNA are also within the scope of the invention. Thus the invention also provides the promoter of an enzyme as defined above.

In a preferred embodiment of this aspect of the  
25 invention, there is provided a nucleic acid construct comprising a nucleic acid sequence of the invention, a biologically-active fragment thereof, or a fragment thereof encoding a biologically-active fragment of an enzyme as defined above, operably linked to one or more nucleic acid  
30 sequences facilitating expression of said enzyme in a plant, preferably a cereal plant. The construct may be a plasmid or a vector, preferably one suitable for use in the transformation of a plant. A particularly suitable vector is a bacterium of the genus *Agrobacterium*, preferably  
35 *Agrobacterium tumefaciens*. Methods of transforming cereal plants using *Agrobacterium tumefaciens* are known; see for example Australian Patent No. 667939 by Japan Tobacco Inc.,

- 116 -

## CLAIMS

1. A nucleic acid sequence encoding an enzyme of the starch biosynthetic pathway in a cereal plant, wherein the enzyme is selected from the group consisting of starch branching enzyme I, starch branching enzyme II, starch soluble synthase I, and debranching enzyme, with the proviso that the enzyme is not soluble starch synthase I of rice, or starch branching enzyme I of rice or maize.
2. A sequence according to claim 1, wherein the sequence is a genomic DNA or cDNA sequence.
3. A sequence according to claim 1 or claim 2, wherein the sequence is functional in wheat.
4. A sequence according to any one of claims 1 to 3, wherein the sequence is derived from a *Triticum* species.
5. A sequence according to claim 4, wherein the *Triticum* species is *Triticum tauschii*.
6. A sequence according to any one of claims 1 to 5, wherein the sequence encodes starch branching enzyme I or a biologically-active fragment thereof, and wherein the sequence has at least 70% sequence homology with the sequence shown in SEQ ID NO:5 or SEQ ID NO:9.
7. A sequence according to claim 6, wherein the homology is at least 90%.
8. A sequence according to any one of claims 1 to 5, wherein the sequence encodes starch branching enzyme II or a biologically-active fragment thereof, and wherein the sequence has at least 70% sequence homology with the sequence shown in SEQ ID NO:10.

- 117 -

9. A sequence according to claim 8, wherein the homology is at least 90%.
10. A sequence according to any one of claims 1 to 5,  
5 wherein the sequence encodes soluble starch synthase or a biologically-active fragment thereof, and wherein the sequence has at least 70% sequence homology with the sequence shown in SEQ ID NO:11 or SEQ ID NO:13.
- 10 11. A sequence according to claim 10, wherein the homology is at least 90%.
12. A sequence according to claim 11, wherein the sequence encodes a 75 kD soluble starch synthase of wheat.  
15
13. A sequence according to claim 12, which encodes an amino acid sequence at least 70% homologous to that shown in SEQ ID NO:14.
- 20 14. A sequence according to any one of claims 1 to 5, wherein the sequence encodes debranching enzyme or a biologically-active fragment thereof, and wherein the sequence has at least 70% sequence homology with the sequence shown in SEQ ID No:17.
- 25 15. A sequence according to claim 14, wherein the homology is at least 90%.
16. A promoter of an enzyme selected from the group  
30 consisting of starch branching enzyme I, starch branching enzyme II, starch soluble synthase I, and debranching enzyme, with the proviso that the enzyme is not soluble starch synthase I of rice, or starch branching enzyme I of rice or maize.
- 35 17. A promoter according to claim 16, wherein the promoter is a starch branching enzyme I promoter or

- 118 -

biologically-active fragment thereof, and wherein the promoter sequence has at least 70% sequence homology with the sequence shown in SEQ ID No:8.

5 18. A sequence according to claim 17, wherein the homology is at least 90%.

19. A promoter according to claim 16, wherein the promoter is a starch soluble synthase I promoter or  
10 biologically-active fragment thereof, and wherein the promoter sequence has at least 70% sequence homology with the sequence shown in SEQ ID No:15.

20. A sequence according to claim 19, wherein the  
15 homology is at least 90%.

21. A nucleic acid construct comprising a nucleic acid sequence encoding an enzyme of the starch biosynthetic pathway in a cereal plant, operably linked to one or more  
20 nucleic acid sequences facilitating expression of the nucleic acid sequence in a plant, wherein the enzyme is selected from the group consisting of starch branching enzyme I, starch branching enzyme II, starch soluble synthase I, and debranching enzyme, with the proviso that  
25 the enzyme is not soluble starch synthase I of rice, or starch branching enzyme I of rice or maize, a biologically-active fragment thereof.

22. A nucleic acid construct for targeting a gene to  
30 the endosperm of a cereal plant, comprising one or more promoter sequences selected from the group consisting of SBE I promoter, SBE II promoter, SSS I promoter, and DBE promoter, operatively linked to a nucleic acid sequence encoding a protein, wherein the expression of the targetted  
35 gene in the endosperm of a cereal plant is modified.



- 119 -

23. A construct according to either claim 21 or claim 22, wherein the promoter or nucleic acid sequence is also operatively linked to one or more additional targeting sequences and/or one or more 3' untranslated sequences.
- 5 24. A construct according to claim 23, wherein the nucleic acid encoding the protein is either in the sense or antisense orientation.
- 10 25. A construct according to claims 24, wherein the protein is an enzyme of the starch biosynthetic pathway.
26. A construct according to claim 25, wherein the nucleic acid encoding the protein is in the antisense  
15 orientation, and the enzyme is selected from the group consisting of GBSS, starch debranching enzyme, SBE II, low molecular weight glutenin, and grain softness protein I.
27. A construct according to claim 25, wherein the  
20 nucleic acid encoding the protein is in the sense orientation, and the enzyme is selected from the group consisting of bacterial isoamylase, bacterial glycogen synthase, and wheat high molecular weight glutenin Bx17.
28. A construct according to any one of claims 21 to  
25 27, wherein the plant is a cereal plant.
29. A construct according to claim 28, wherein the cereal plant is either wheat or barley.
- 30 30. A construct according to claim 29, wherein the cereal plant is wheat.
31. A construct according to any one of claims 21 to  
30, wherein the construct is either a plasmid or a vector.
- 35

- 120 -

32. A construct according to claim 31, wherein the plasmid or vector is suitable for use in the transformation of a plant.
- 5 33. A construct according to claim 32, wherein the plasmid is selected from the group consisting of those depicted in Figures 22a to 22f.
34. A construct according to claim 32, wherein the  
10 vector is a bacterium of the genus *Agrobacterium*.
35. A construct according to claim 34, wherein the vector is *Agrobacterium tumefaciens*.
- 15 36. A method of modifying the characteristics of starch produced by a plant, comprising the steps of:  
(a) introducing a nucleic acid sequence encoding an enzyme of the starch biosynthetic pathway into a host plant, and/or  
20 (b) introducing an anti-sense nucleic acid sequence directed to a gene encoding an enzyme of the starch biosynthetic pathway into a host plant,  
wherein the enzyme is selected from the group consisting of starch branching enzyme I, starch branching  
25 enzyme II, starch soluble synthase I, and debranching enzyme, with the proviso that the enzyme is not soluble starch synthase I of rice, or starch branching enzyme I of rice or maize, and wherein if both steps (a) and (b) are used, the enzymes in the two steps are different.  
30
37. A method according to claim 36, wherein the plant is a cereal plant.
38. A method according to claim 37, wherein the cereal  
35 plant is wheat or barley.

- 121 -

39. A method of targeting expression of a gene to the endosperm of a cereal plant, comprising the step of transforming the plant with a construct according to any one of claims 21 to 35.

5

40. A method of modulating the time of expression of a gene in endosperm of a cereal plant, comprising the step of transforming the plant with a construct according to any one of claims 21 to 35.

10

41. A method according to claim 40, wherein when expression at an early stage following anthesis is desired, the construct comprises either the SBE II, SSS I, or DBE promoter.

15

42. A method according to claim 40, wherein when expression at a later stage following anthesis is desired, the construct comprises the SBE I promoter.

20

43. A plant transformed with a construct according to any one of claims 21 to 35.

44. A plant according to claim 43, wherein the plant is a cereal plant.

25

45. A plant according to claim 44, wherein the cereal plant is wheat or barley.

46. A method of identifying variations in the starch synthesis characteristics of a cereal plant, comprising the step of identifying a variation in nucleic acid sequence in the intron regions of the SBE I, SBE II, SSS I or DBE genes.

30

47. A method of identifying variations in the starch synthesis characteristics of a cereal plant, comprising the step of identifying a variation in nucleic acid sequence compared to the sequence shown in one or more SEQ ID NO:5,

35

- 122 -

SEQ ID NO:7, SEQ ID NO:9, SEQ ID NO:10, SEQ ID NO:11, SEQ ID NO:13, SEQ ID NO:15, SEQ ID NO:16, or SEQ ID NO:17.

48. A method according to claim 47, in which a mutation or absence of a SBE I, SBE II, SSS I or DBE gene is detected.

49. A method according to either claim 47 or claim 48, in which the cereal plant is wheat or barley.

50. A product comprising plant material propagated from a plant transformed with a nucleic acid sequence encoding an enzyme of the starch biosynthetic pathway in a cereal plant, operably linked to one or more nucleic acid sequences facilitating expression of the nucleic acid sequence in a plant, wherein the enzyme is selected from the group consisting of starch branching enzyme I, starch branching enzyme II, starch soluble synthase I, and debranching enzyme, with the proviso that the enzyme is not soluble starch synthase I of rice, or starch branching enzyme I of rice or maize, a biologically-active fragment thereof.

51. A product comprising plant material propagated from a plant in which a gene was targeted to the endosperm of a cereal plant, by a nucleic acid construct comprising one or more promoter sequences selected from the group consisting of SBE I promoter, SBE II promoter, SSS I promoter, and DBE promoter, operatively linked to a nucleic acid sequence encoding a protein, wherein the expression of the targetted gene in the endosperm of a cereal plant is modified.

52. A product according to claim 50 or claim 51 wherein the product is a food product.

**PCT**WORLD INTELLECTUAL PROPERTY ORGANIZATION  
International Bureau

## INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

<b>(51) International Patent Classification <sup>6</sup> :</b> <b>C12N 9/24, 15/55</b>	<b>A1</b>	<b>(11) International Publication Number:</b> <b>WO 99/14314</b> <b>(43) International Publication Date:</b> 25 March 1999 (25.03.99)
<b>(21) International Application Number:</b> PCT/AU98/00743 <b>(22) International Filing Date:</b> 11 September 1998 (11.09.98)  <b>(30) Priority Data:</b> PO 9108 ✓ 12 September 1997 (12.09.97) AU PP 2509 ✓ 20 March 1998 (20.03.98) AU  <b>(71) Applicants (for all designated States except US):</b> COMMON-WEALTH SCIENTIFIC AND INDUSTRIAL RESEARCH ORGANISATION [AU/AU]; Limestone Avenue, Campbell, ACT 2612 (AU). THE AUSTRALIAN NATIONAL UNIVERSITY [AU/AU]; Acton, ACT 2601 (AU). GOODMAN FIELDER LIMITED [AU/AU]; Level 42, Grosvenor Place, Sydney, NSW 2000 (AU). GROUPE LIMAGRAIN PACIFIC PTY. LIMITED [AU/AU]; Level 31, 1 O'Connell Street, Sydney, NSW 2000 (AU).  <b>(72) Inventors; and</b> <b>(75) Inventors/Applicants (for US only):</b> LI, Zhongyi [CN/AU]; 63 Campaspe Circuit, Kaleen, ACT 2617 (AU). MORELL, Matthew [AU/AU]; 33 Wangara Street, Aranda, ACT 2614 (AU). RAHMAN, Sadequr [AU/AU]; 46 Scarlett Street, Melba, ACT 2615 (AU).		<b>(74) Agent:</b> GRIFFITH HACK; 509 St. Kilda Road, Melbourne, VIC 3004 (AU).  <b>(81) Designated States:</b> AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GE, GH, GM, HR, HU, ID, IL, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, US, UZ, VN, YU, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).  <b>Published</b> With international search report.
<b>(54) Title:</b> REGULATION OF GENE EXPRESSION IN PLANTS		
<b>(57) Abstract</b>  The present invention relates to a nucleic acid sequence encoding an enzyme of the starch biosynthetic pathway in a cereal plant, wherein the enzyme is selected from the group consisting of starch branching enzyme I, starch branching enzyme II, starch soluble synthase I, and debranching enzyme, with the proviso that the enzyme is not soluble starch synthase I of rice, or starch branching enzyme I of rice or maize.		

**FOR THE PURPOSES OF INFORMATION ONLY**

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	ES	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
AT	Austria	FR	France	LU	Luxembourg	SN	Senegal
AU	Australia	GA	Gabon	LV	Larvia	SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Monaco	TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistan
BE	Belgium	GN	Guinea	MK	The former Yugoslav Republic of Macedonia	TM	Turkmenistan
BF	Burkina Faso	GR	Greece	ML	Mali	TR	Turkey
BG	Bulgaria	HU	Hungary	MN	Mongolia	TT	Trinidad and Tobago
BJ	Benin	IE	Ireland	MR	Mauritania	UA	Ukraine
BR	Brazil	IL	Israel	MW	Malawi	UG	Uganda
BY	Belarus	IS	Iceland	MX	Mexico	US	United States of America
CA	Canada	IT	Italy	NE	Niger	UZ	Uzbekistan
CF	Central African Republic	JP	Japan	NL	Netherlands	VN	Viet Nam
CG	Congo	KE	Kenya	NO	Norway	YU	Yugoslavia
CH	Switzerland	KG	Kyrgyzstan	NZ	New Zealand	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	Democratic People's Republic of Korea	PL	Poland		
CM	Cameroon	KR	Republic of Korea	PT	Portugal		
CN	China	KZ	Kazakstan	RO	Romania		
CU	Cuba	LC	Saint Lucia	RU	Russian Federation		
CZ	Czech Republic	LI	Liechtenstein	SD	Sudan		
DE	Germany	LK	Sri Lanka	SE	Sweden		
DK	Denmark	LR	Liberia	SG	Singapore		
EE	Estonia						

- 1 -

## REGULATION OF GENE EXPRESSION IN PLANTS

This invention relates to methods of modulating the expression of desired genes in plants, and to DNA sequences and genetic constructs for use in these methods. In particular, the invention relates to methods and constructs for targeting of expression specifically to the endosperm of the seeds of cereal plants such as wheat, and for modulating the time of expression in the target tissue. This is achieved by the use of promoter sequences from enzymes of the starch biosynthetic pathway. In a preferred embodiment of the invention, the sequences and/or promoters are those of starch branching enzyme I, starch branching enzyme II, soluble starch synthase I, and starch debranching enzyme, all derived from *Triticum tauschii*, the D genome donor of hexaploid bread wheat.

A further preferred embodiment relates to a method of identifying variations in the characteristics of plants.

## 20 BACKGROUND OF THE INVENTION

Starch is an important constituent of cereal grains and of flours, accounting for about 65-67% of the weight of the grain at maturity. It is produced in the amyloplast of the grain endosperm by the concerted action of a number of enzymes, including ADP-Glucose pyrophosphorylase (EC 2.7.7.27), starch synthases (EC 2.4.1.21), branching enzymes (EC 2.4.1.18) and debranching enzymes (EC 3.2.1.41 and EC 3.2.1.68) (Ball et al, 1996; Martin and Smith, 1995; Morell et al, 1995). Some of the proteins involved in the synthesis of starch can be recovered from the starch granule (Denyer et al, 1995; Rahman et al, 1995).

Most wheat cultivars normally produce starch containing 25% amylose and 75% amylopectin. Amylose is composed of large linear chains of  $\alpha$  (1-4) linked  $\alpha$ -D-glucopyranosyl residues, whereas amylopectin is a branching form of  $\alpha$ -glycan linked by  $\alpha$  (1-6) linkages. The ratio of amylose and amylopectin, the branch chain length and the

- 2 -

number of branch chains of amylopectin are the major factors which determine the properties of wheat starch.

Starch with various properties has been widely used in industry, food science and medical science. High amylose wheat can be used for plastic substitutes and in paper manufacture to protect the environment; in health foods to reduce bowel cancer and heart disease; and in sports foods to improve the athletes' performance. High amylopectin wheat may be suitable for Japanese noodles, and is used as a thickener in the food industry.

Wheat contains three sets of chromosomes (A, B and D) in its very large genome of about  $10^{10}$  base pairs (bp). The donor of the D genome to wheat is *Triticum tauschii*, and by using a suitable accession of this species the genes from the D genome can be studied separately (Lagudah et al, 1991).

There is comparatively little variation in starch structure found in wheat varieties, because the hexaploid nature of wheat prevents mutations from being readily identified. Dramatic alterations in starch structure are expected to require the combination of homozygous recessive alleles from each of the 3 wheat genomes, A, B and D. This requirement renders the probability of finding such mutants in natural or mutagenised populations of wheat very low. Variation in wheat starch is desirable in order to enable better tailoring of wheat starches for processing and end-user requirements.

Key commercial targets for the manipulation of starch biosynthesis are:

1. "Waxy" wheats in which amylose content is decreased to insignificant levels. This outcome is expected to be obtained by eliminating granule-bound starch synthase activity.
2. High amylose wheats, expected to be obtained by suppressing starch branching enzyme-II activity.
3. Wheats which continue to synthesise starch at elevated temperatures, expected to be obtained by



- 3 -

identifying or introducing a gene encoding a heat-stable soluble starch synthase.

4. "Sugary types" of wheat which contain increased amylose content and free sugars, expected to be obtained by manipulating an isoamylase-type debranching enzyme.

There are two general strategies which may be used to obtain wheats with altered starch structure:

- (a) using genetic engineering strategies to suppress the activity of a specific gene, or to introduce a novel gene into a wheat line; and
- (b) selecting among existing variation in wheat for missing ("null") or altered alleles of a gene in each of the genomes of wheat, and combining these by plant breeding.

However, in view of the complexity of the gene families, particularly starch branching enzyme I (SBE I), without the ability to target regions which are unique to genes expressed in endosperm, modification of wheat by combination of null alleles of several enzymes in general represents an almost impossible task.

Branching enzymes are involved in the production of glucose  $\alpha$ -1,6 branches. Of the two main constituents of starch, amylose is essentially linear, but amylopectin is highly branched; thus branching enzymes are thought to be directly involved in the synthesis of amylopectin but not amylose. There are two types of branching enzymes in plants, starch branching enzyme I (SBE I) and starch branching enzyme II (SBE II), and both are about 85 kDa in size. At the nucleic acid level there is about 65% sequence identity between types I and II in the central portion of the molecules; the sequence identity between SBE I from different cereals is about 85% overall (Burton et al, 1995; Morell et al, 1995).

In cereals, SBE I genes have so far been reported only for rice (Kawasaki et al, 1991; Rahman et al, 1997). A cDNA sequence for wheat SBE I is available on the GenBank

database (Accession No. Y12320; Repellin A., Nair R.B., Baga M., and Chibbar R.N.: Plant Gene Register PGR97-094, 1997). As far as we are aware, no promoter sequence for wheat SBE I has been reported.

5           We have characterised an SBE I gene, designated *wSBE I-D2*, from *Triticum tauschii*, the donor of the D genome to wheat (Rahman et al, 1997). This gene encoded a protein sequence which had a deletion of approximately 65 amino acids at the C-terminal end, and appeared not to contain  
10 some of the conserved amino acid motifs characteristic of this class of enzyme (Svensson, 1994). Although *wSBE I-D2* was expressed as mRNA, no corresponding protein has yet been found in our analysis of SBE I isoforms from the endosperm, and thus it is possible that this gene is a transcribed  
15 pseudogene.

Genes for SBE II are less well characterised; no genomic sequences are available, although SBE II cDNAs from rice (Mizuno et al, 1993; Accession No. D16201) and maize (Fisher et al, 1993; Accession No. L08065) have been  
20 reported. In addition, a cDNA sequence for SBE II from wheat is available on the GenBank database (Nair et al, 1997; Accession No. Y11282); although the sequences are very similar to those reported herein, there are differences near the N-terminal of the protein, which specifies its  
25 intracellular location. No promoter sequences have been reported, as far as we are aware.

Wheat granule-bound starch synthase (GBSS) is responsible for amylose synthesis, while wheat branching enzymes together with soluble starch synthases are  
30 considered to be directly involved in amylopectin biosynthesis. A number of isoforms of soluble and granule-bound starch synthases have been identified in developing wheat endosperm (Denyer et al, 1995). There are three distinct isoforms of starch synthases, 60 kDa, 75-77 kDa and  
35 100-105 kDa, which exist in the starch granules (Denyer et al, 1995; Rahman et al, 1995). The 60 kDa GBSS is the product of the *wx* gene. The 75-77 kDa protein is a wheat

- 5 -

soluble starch synthase I (SSSI) which is present in both the soluble fraction and the starch granule-bound fraction of the endosperm. However, the 100-105 kDa proteins, which are another type of soluble starch synthase, are located  
5 only in starch granules (Denyer *et al*, 1995; Rahman *et al*, 1995). To our knowledge there has been no report of any complete wheat SSS I sequence, either at the protein or the nucleotide level.

Both cDNA and genomic DNA encoding a soluble  
10 starch synthase I of rice have been cloned and analysed (Baba *et al*, 1993; Tanaka *et al*, 1995). The cDNAs encoding potato soluble starch synthase SSSII and SSSIII and pea soluble starch synthase SSSII have also been reported (Edwards *et al*, 1995; Marshall *et al*, 1996; Dry *et al*,  
15 1992). However, corresponding full length cDNA sequences for wheat have hitherto not been available, although a partial cDNA sequence (Accession No. U48227) has been released to the GenBank database.

Approach (b) referred to above has been  
20 demonstrated for the gene for granule-bound starch synthase. Null alleles on chromosomes 7A, 7D and 4A were identified by the analysis of GBSS protein bands by electrophoresis, and combined by plant breeding to produce a wheat line containing no GBSS, and no amylose (Nakamura *et al*, 1995).  
25 Subsequently, PCR-based DNA markers have been identified, which also identify null alleles for the GBSS loci on each of the three wheat genomes. Despite the availability of a considerable amount of information in the prior art, major problems remain. Firstly, the presence of three separate  
30 sets of chromosomes in wheat makes genetic analysis in this species extraordinarily complex. This is further complicated by the fact that a number of enzymes are involved in starch synthesis, and each of these enzymes is itself present in a number of forms, and in a number of  
35 locations within the plant cell. Little, if any, information has been available as to which specific form of each enzyme is expressed in endosperm. For wheat, a limited

- 6 -

amount of nucleic acid sequence information is available, but this is only cDNA sequence; no genomic sequence, and consequently no information regarding promoters and other control sequences, is available. Without being able to demonstrate that the endosperm-specific gene within a family has been isolated, such sequence information is of limited practical usefulness.

#### SUMMARY OF THE INVENTION

In this application we report the isolation and identification of novel genes from *T. tauschii*, the D-genome donor of wheat, that encode SBE I, SBE II, a 75 kDa SSS I, and an isoamylase-type debranching enzyme (DBE). Because of the very close relationship between *T. tauschii* and wheat, as discussed above, results obtained with *T. tauschii* can be directly applied to wheat with little if any modification. Such modification as may be required represents routine trial and error experimentation. Sequences from these genes can be used as probes to identify null or altered alleles in wheat, which can then be used in plant breeding programmes to provide modifications of starch characteristics. The novel sequences of the invention can be used in genetic engineering strategies or to introduce a desired gene into a host plant, to provide antisense sequences for suppression of one or more specific genes in a host plant, in order to modify the characteristics of starch produced by the plant.

By using *T. tauschii*, we have been able to examine a single genome, rather than three as in wheat, and to identify and isolate the forms of the starch synthesis genes which are expressed in endosperm. By addressing genomic sequences we have been able to isolate tissue-specific promoters for the relevant genes, which provides a mechanism for simultaneous manipulation of a number of genes in the endosperm. Because *T. tauschii* is so closely related to wheat, results obtained with this model system are directly applicable to wheat, and we have confirmed this experimentally. The genomic sequences which we have

- 7 -

determined can also be used as probes for the identification and isolation of corresponding sequences, including promoter sequences, from other cereal plant species.

In its most general aspect, the invention provides  
5 a nucleic acid sequence encoding an enzyme of the starch biosynthetic pathway in a cereal plant, said enzyme being selected from the group consisting of starch branching enzyme I, starch branching enzyme II, starch soluble synthase I, and debranching enzyme, with the proviso that  
10 the enzyme is not soluble starch synthase I of rice, or starch branching enzyme I of rice or maize.

Preferably the nucleic acid sequence is a DNA sequence, and may be genomic DNA or cDNA. Preferably the sequence is one which is functional in wheat. More  
15 preferably the sequence is derived from a *Triticum* species, most preferably *Triticum tauschii*.

Where the sequence encodes soluble starch synthase, preferably the sequence encodes the 75 kD soluble starch synthase of wheat.

20 Biologically-active untranslated control sequences of genomic DNA are also within the scope of the invention. Thus the invention also provides the promoter of an enzyme as defined above.

In a preferred embodiment of this aspect of the  
25 invention, there is provided a nucleic acid construct comprising a nucleic acid sequence of the invention, a biologically-active fragment thereof, or a fragment thereof encoding a biologically-active fragment of an enzyme as defined above, operably linked to one or more nucleic acid  
30 sequences facilitating expression of said enzyme in a plant, preferably a cereal plant. The construct may be a plasmid or a vector, preferably one suitable for use in the transformation of a plant. A particularly suitable vector is a bacterium of the genus *Agrobacterium*, preferably  
35 *Agrobacterium tumefaciens*. Methods of transforming cereal plants using *Agrobacterium tumefaciens* are known; see for example Australian Patent No. 667939 by Japan Tobacco Inc.,

International Patent Application Number PCT/US97/10621 by Monsanto Company and Tingay et al (1997).

5 In a second aspect, the invention provides a nucleic acid construct for targeting of a desired gene to endosperm of a cereal plant, and/or for modulating the time of expression of a desired gene in endosperm of a cereal plant, comprising one or more promoter sequences selected from SBE I promoter, SBE II promoter, SSS I promoter, and DBE promoter, operatively linked to a nucleic acid sequence  
10 encoding a desired protein, and optionally also operatively linked to one or more additional targeting sequences and/or one or more 3' untranslated sequences.

The nucleic acid encoding the desired protein may be in either the sense orientation or in the antisense  
15 orientation. Preferably the desired protein is an enzyme of the starch biosynthetic pathway. For example, the antisense sequences of GBSS, starch debranching enzyme, SBE II, low molecular weight glutenin, or grain softness protein I, may be used. Preferred sequences for use in sense orientation  
20 include those of bacterial isoamylase, bacterial glycogen synthase, or wheat high molecular weight glutenin Bx17. It is contemplated that any desired protein which is encoded by a gene which is capable of being expressed in the endosperm of a cereal plant is suitable for use in the invention.

25 In a third aspect, the invention provides a method of modifying the characteristics of starch produced by a plant, comprising the step of:

- (a) introducing a gene encoding a desired enzyme of the starch biosynthetic pathway into a host plant, and/or  
30 (b) introducing an anti-sense nucleic acid sequence directed to a gene encoding an enzyme of the starch biosynthetic pathway into a host plant,

wherein said enzymes are as defined above.

Where both steps (a) and (b) are used, the enzymes  
35 in the two steps are different.

Preferably the plant is a cereal plant, more preferably wheat or barley.

- 9 -

As is well known in the art, anti-sense sequences can be used to suppress expression of the protein to which the anti-sense sequence is complementary. It will be evident to the person skilled in the art that different  
5 combinations of sense and anti-sense sequences may be chosen so as to effect a variety of different modifications of the characteristics of the starch produced by the plant.

In a fourth aspect, the invention provides a method of targeting expression of a desired gene to the  
10 endosperm of a cereal plant, comprising the step of transforming the plant with a construct according to the invention.

According to a fifth aspect, the invention provides a method of modulating the time of expression of a  
15 desired gene in endosperm of a cereal plant, comprising the step of transforming the plant with a construct according to the second aspect of the invention.

Where expression at an early stage following anthesis is desired, the construct preferably comprises the  
20 SBE II, SSS I or DBE promoters. Where expression at a later stage following anthesis is desired, the construct preferably comprises the SBE I promoter.

While the invention is described in detail in relation to wheat, it will be clearly understood that it is  
25 also applicable to other cereal plants of the family Gramineae, such as maize, barley and rice.

Methods for transformation of monocotyledonous plants such as wheat, maize, barley and rice and for  
regeneration of plants from protoplasts or immature plant  
30 embryos are well known in the art. See for example Lazzeri et al, 1991; Jahne et al, 1991 and Wan and Lemaux, 1994 for barley; Wirtzens et al, 1997; Tingay et al, 1997; Canadian Patent Application No. 2092588 by Nehra; Australian Patent Application No. 61781/94 by National Research Council of  
35 Canada, Australian Patent No. 667939 by Japan Tobacco Co, and International Patent Application Number PCT/US97/10621 by Monsanto Company.

- 10 -

The sequences of ADP glucose pyrophosphorylase from barley (Australian Patent Application No. 65392/94), starch debranching enzyme and its promoter from rice (Japanese Patent Publication No. Kokai 6261787 and Japanese Patent Publication No. Kokai 5317057), and starch debranching enzyme from spinach and potato (Australian Patent Application No. 44333/96) are all known.

#### Detailed Description of the Drawings

The invention will be described in detail by reference only to the following non-limiting examples and to the figures.

Figure 1 shows the hybridisation of genomic clones isolated from *T. tauschii*.

DNA was extracted from the different clones, digested with *Bam*HI and hybridised with the 5' end of the maize SBE I cDNA. Lanes 1, 2, 3 and 4 correspond to DNA from clones  $\lambda$ E1,  $\lambda$ E2,  $\lambda$ E6 and  $\lambda$ E7 respectively. Note that clones  $\lambda$ E1 and  $\lambda$ E2 give identical patterns, the SBE I gene in  $\lambda$ E6 is a truncated form of that in  $\lambda$ E1, and  $\lambda$ E7 gives a clearly different pattern.

Figure 2 shows the hybridisation of DNA from *T. tauschii*.

DNA from *T. tauschii* was digested with *Bam*HI and the hybridisation pattern compared with DNA from  $\lambda$ E1 and  $\lambda$ E7 digested with the same enzyme. Fragment E1.1 (see Figure 3) from  $\lambda$ E1 was used as the probe; it contains some sequences that are over 80% identical to sequences in E7.8. Approximately 25  $\mu$ g of *T. tauschii* DNA was electrophoresed in lane 1, and 200 pg each of  $\lambda$ E1 and  $\lambda$ E7 in lanes 2 and 3, respectively.

Figure 3 shows the restriction maps of clone  $\lambda$ E1 and  $\lambda$ E7. The fragments obtained with *Eco*RI and *Bam*HI are indicated. The fragments sequenced from  $\lambda$ E1 are E1.1, E1.2, a part of E1.7 and a part of E1.5.

Figure 4 shows the comparison of deduced amino acid sequence of wSBE I-D4 cDNA with the deduced amino acid



- 11 -

sequence of rice SBE I (RSBE I; Nakamura et al, 1992), maize SBE I (MSBE I; Baba et al, 1991), wSBE I-D2 type cDNA (D2 cDNA; Rahman et al, 1997), pea SBE II (PESBE II, homologous to maize SBE I; Burton et al, 1995), and potato SBE I (POSBE; Cangiano et al, 1993). The deduced amino acid sequence of the wSBE I-D4 cDNA is denoted by "D4cDNA". Residues present in at least three of the sequences are identified in the consensus sequence in capitals.

Figure 5 shows the intron-exon structure of wSBE I-D4 compared to the corresponding structures of rice SBE I (Kawasaki et al, 1993) and wSBE I-D2 (Rahman et al, 1997). The intron-exon structure of wSBE I-D4 is deduced by comparison with the SBE I cDNA reported by Repellin et al (1997).

The dark rectangles correspond to exons and the light rectangles correspond to introns. The bars above the structures indicate the percentage identity in sequence between the indicated exons and introns of the relevant genes. Note that intron 2 shares no significant sequence identity and is not indicated.

Figure 6 shows the nucleotide sequence of part of wSBE I-D4, the amino acid sequence deduced from this nucleotide sequence, and the N-terminal amino acid sequence of the SBE I purified from the wheat endosperm (Morell et al, 1997).

Figure 7 shows the hybridisation of SBE I genomic clones with the following probes,

A. wSBE I-D45 (derived from the 5' end of the gene and including sequence from fragments E1.1 and E1.7), and

B. wSBE I-D43 (derived from the 3' end of the gene and containing sequences from fragment E1.5). For panel A, the tracks 1-13 correspond to clones  $\lambda$ E1,  $\lambda$ E2,  $\lambda$ E6,  $\lambda$ E7,  $\lambda$ E9,  $\lambda$ E14,  $\lambda$ E22,  $\lambda$ E27, Molecular weight markers,  $\lambda$ E29,  $\lambda$ E30,  $\lambda$ E31 and  $\lambda$ E52. For panel B, tracks 1-12 correspond to clones  $\lambda$ E1,  $\lambda$ E2,  $\lambda$ E6,  $\lambda$ E7,  $\lambda$ E9,  $\lambda$ E14,  $\lambda$ E22,  $\lambda$ E27,  $\lambda$ E29,  $\lambda$ E30,  $\lambda$ E31 and  $\lambda$ E52. Note that clones  $\lambda$ E7 and  $\lambda$ E22 do not

- 12 -

hybridise to either of the probes and are wSBE I-D2 type genes. Also note that clone  $\lambda$ E30 contains a sequence unrelated to SBE I. The size of the molecular weight markers in kb is indicated. Clones  $\lambda$ E7 and  $\lambda$ E22 do  
5 hybridise with a probe from E1.1. which is highly conserved between wSBE I-D2 and wSBE I-D4.

Figure 8 shows the alignment of cDNA clones to obtain the sequence represented by wSBE I-D4 cDNA. BED4 and BED5 were obtained from screening the cDNA library with  
10 maize BEI (Baba et al, 1991). BED1, 2 and 3 were obtained by RT-PCR using defined primers.

Figure 9a shows the expression of Soluble Starch Synthase I (SSS), Starch Branching Enzyme I (BE I) and Starch Branching Enzyme II (BE II) mRNAs during endosperm  
15 development.

RNA was purified from leaves, florets prior to anthesis, and endosperm of wheat cultivar Rosella grown in a glasshouse, collected 5 to 8 days after anthesis, 10 to 15 days after anthesis and 18 to 22 days after anthesis, and  
20 from the endosperm of wheat cultivar Rosella grown in the field and collected 12, 15 and 18 days after anthesis respectively. Equivalent amounts of RNA were electrophoresed in each lane. The probes were from the coding region of the SM2 SSS I cDNA (from nucleotide 1615 to  
25 1919 of the SM2 cDNA sequence); wSBE I-D43C (see Table I), which corresponds to the untranslated 3' end of wSBE I-D4 cDNA (E1 (3'; and the 5' region of SBE9 (SBE9 (5'), corresponding to the region between nucleotides 743 to 1004 of Genbank sequence Y11282. No hybridisation to RNA  
30 extracted from leaves or preanthesis florets was detected.

Figure 9b shows the hybridisation of RNA from the endosperm of the hexaploid *T. aestivum* cultivar "Gabo" with the starch branching enzyme I gene. The probe, wSBEI-D43, is defined in Table 1.

35 Figure 9c shows the hybridisation of RNA from the endosperm of the hexaploid *T. aestivum* cultivar "Wyuna" with

- 13 -

the starch branching enzyme II gene. The probe, wSBE II-D13, is defined in Table 2.

Figure 9d shows the hybridisation of RNA from the endosperm of the hexaploid *T. aestivum* cultivar "Gabo" with the SSS I gene. The probe spanned the region from nucleotides 2025 to 2497 of the SM2 cDNA sequence shown in SEQ ID No:11.

Figure 9e shows the hybridisation of RNA from the endosperm of the hexaploid *T. aestivum* cultivar "Gabo" with the DBE I gene. The probe, a DBE3' 3'PCR fragment, extends from nucleotide position 281 to 1072 of the cDNA sequence in SEQ ID No:16.

Figure 9f shows the hybridisation of RNA from the endosperm of the hexaploid *T. aestivum* cultivar "Gabo" with the wheat actin gene. The probe was a wheat actin DNA sequence generated by PCR from wheat endosperm cDNA using primers to conserved plant actin sequences.

Figure 9g shows the hybridisation of RNA from the endosperm of the hexaploid *T. aestivum* cultivar "Gabo" with a probe containing wheat ribosomal RNA 26S and 18S fragments (plasmid pta250.2 from Dr Bryan Clarke, CSIRO Plant Industry).

Figure 9h shows the hybridisation of RNA from the hexaploid wheat cultivar "Gabo" with the DBE I probe described in Figure 9e. Lane 1; leaf RNA; lane 2, pre-anthesis floret RNA; lane 3, RNA from endosperm harvested 12 days after anthesis.

Figure 10 shows the comparison of wSBE I-D4 (sr 427.res ck: 6,362,1 to 11,099) and rice SBE I genomic sequence (d10838.em\_pl ck: 3,071,1 to 11,700) (Kawasaki et al, 1993; Accession Number D10838) using the programs Compares and DotPlot (Devereaux et al, 1984). The programs used a window of 21 bases with a stringency of 14 to register a dot.

Figure 11 shows the hybridisation of wheat DNA from chromosome-engineered lines using the following probes:

A. wSBE I-D45 (from the 5' end of the gene),

- 14 -

B. wSBE I-D43 (from the 3' end of the gene),  
and

C. wSBE I-D4R (repetitive sequence  
approximately 600 bp 3' to the end of wSBE I-D4 sequence.

5 N7AT7B, no 7A chromosome, four copies of 7B  
chromosome; N7BT7D, no 7B chromosome, four copies of 7D  
chromosome; NTDT7A, no 7D chromosome, four copies of 7A  
chromosome. The chromosomal origin of hybridising bands is  
indicated.

10 Figure 12 shows the hybridisation of genomic  
clones F1, F2, F3 and F4 with the entire SBE-9 sequence.  
The DNA from the clones was purified and digested with  
either *Bam*HI or *Eco*RI, separated on agarose, blotted onto  
nitrocellulose and hybridised with labelled SBE-9 (a SBE II  
15 type cDNA). The pattern of hybridising bands is different  
in the four isolates.

Figure 13a shows the N-terminal sequence of  
purified SBE II from wheat endosperm as in Morell *et al*,  
(1997).

20 Figure 13b shows the deduced amino acid sequence  
from part of wSBE II-D1 that encodes the N-terminal sequence  
as described in Morell *et al*, (1997).

Figure 14 shows the deduced exon-intron structure  
for a part of wSBE II-D1. The scale is marked in bases.  
25 The dark rectangles are exons.

Figure 15 shows the hybridisation of DNA from  
chromosome engineered lines of wheat (cultivar Chinese  
Spring) with a probe from nucleotides 550-850 from SBE-9.  
The band of approximately 2.2 kb is missing in the line in  
30 which chromosome 2D is absent.

T2BN2A: four copies of chromosome 2B, no copies  
of chromosome 2A;

T2AN2B: four copies of chromosome 2A, no copies  
of chromosome 2B;

35 T2AN2D: four copies of chromosome 2A, no copies  
of chromosome 2D.

- 15 -

Figure 16 shows the N-terminal sequence of SSS I protein isolated from starch granules (Rahman *et al*, 1995) and deduced amino acid sequence of part of Sm2.

Figure 17 shows the hybridisation of genomic clones sgl, 3, 4, 6 and 11 with the cDNA clone (sm2) for SSS I. DNA was purified from indicated genomic clones, digested with *Bam*HI or *Sac*I and hybridised to sm2. Note that the hybridisation patterns for sgl, 3 and 4 are clearly different from each other.

Figure 18 shows a comparison of the intron/exon structures of the wheat and rice soluble starch synthase genomic sequences. The dark rectangles indicate exons and the light rectangles represent introns.

Figure 19 shows the hybridisation of DNA from chromosome engineered lines of wheat (cultivar Chinese Spring) digested with *Pvu*II, with the sm2 probe.

N7AT7B: no 7A chromosome, four copies of 7B chromosome;

N7BT7D: no 7B chromosome, four copies of 7D chromosome;

N7DT7A: no 7D chromosome, four copies of 7A chromosome.

A band is missing in the N7BT7A line.

Figure 20a shows the DNA sequence of a portion of the wheat debranching enzyme (WDBE-1) PCR product. The PCR product was generated from wheat genomic DNA (cultivar Rosella) using primers based on sequences conserved in debranching enzymes from maize and rice.

Figure 20b shows a comparison of the nucleotide sequence of wheat debranching enzyme I (WDBE-I) PCR fragment (WHEAT.DNA) with the maize *Sugary-1* sequence (SUGARY.DNA).

Figure 20c shows a comparison between the intron/exon structures of wheat debranching enzyme gene and the maize *sugary-1* debranching enzyme gene.

Figure 21a shows the results of Southern blotting of *T. tauschii* DNA with wheat DBE-I PCR product. DNA from *T. tauschii* was digested with *Bam*HI, electrophoresed,

- 16 -

blotted and hybridised to the wheat DBE-I PCR product described in Figure 20a. A band of approximately 2 kb hybridised.

Figure 21b shows Chinese Spring nullisomic/  
5 tetrasomic lines probed with probes from the DBE gene. Panel (I) shows hybridisation with a fragment spanning the region from nucleotide 270 to 465 of the cDNA sequence shown in SEQ ID No:16 from the central region of the DBE gene. Panel  
10 (II) shows hybridisation with a probe from the 3' region of the gene, from nucleotide 281 to 1072 of the cDNA sequence given in SEQ ID No:16.

Figures 22a to 22e show diagrammatic  
representations of the DNA vectors used for transient  
expression analysis. In each of the sequences the N-terminal  
15 methionine encoding ATG codon is shown in bold.

Figure 22a shows a DNA construct pwsssIpro1gfpNOT  
containing a 1042 base pair region of the wheat soluble  
starch synthase I promoter (wSSSIpro1, from -1042 to -1, SEQ  
ID No:18) fused to the green fluorescent protein (GFP)  
20 reporter gene.

Figure 22b shows a DNA construct pwsssIpro2gfpNOT  
containing a 3914 base pair region of the wheat soluble  
starch synthase I promoter (wSSSIpro2, from -3914 to -1, SEQ  
ID No:18) fused to the green fluorescent protein (GFP)  
25 reporter gene.

Figure 22c shows a DNA construct psbeIIpro1gfpNOT  
containing an 1203 base pair region of the wheat starch  
branching enzyme II promoter (sbeIIpro1, from 1 to 1023 SEQ  
ID No:10 fused to the green fluorescent protein (GFP)  
30 reporter gene.

Figure 22d shows a DNA construct psbeIIpro2gfpNOT  
containing a 1353 base pair region of the wheat starch  
branching enzyme II promoter and transit peptide coding  
region (sbeIIpro2, regions 1-1203, 1204 to 1336 and 1664 to  
35 1680 of SEQ ID No:10 fused to the green fluorescent protein (GFP) reporter gene.

Figure 22e shows a DNA construct pact\_jsgfg\_nos

- 17 -

containing the plasmid backbone of pSP72 (Promega), the rice *Act1* actin promoter (McElroy et al. 1991), the GFP gene (Sheen et al. 1995) and the *Agrobacterium tumefaciens* nopaline synthase (nos) terminator (Bevan et al. 1983).

5           Figure 23 shows T DNA constructs for stable transformation of rice by *Agrobacterium*. The backbone for each plasmid is p35SH-iC (Wang et al 1997). The various promoter-GFP-Nos regions inserted are shown in (a), (b), (c) and (d) respectively, and are described in detail in Example 10 24. Each of these constructs was inserted into the NotI site of p35SH-iC using the NotI flanking sites at each end of the promoter-GFP-Nos regions. The constructs were named (a) p35SH-iC-BEIIpro1\_GFP\_Nos, (b) p35SH-iC-BEIIpro2\_GFP\_Nos (c) p35SH-iC-SSIpro1\_GFP\_Nos and (d) p35SH-iC- 15 SSIpro2\_GFP\_Nos

Figure 24 illustrates the design of 15 intron-spanning BE II primer sets. Primers were based on wSBE II-D1 sequence (SEQ ID No:10), and were designed such that intron sequences in the wSBE II-D1 sequence (deduced 20 from Figure 13b and Nair et al, 1997; Accession No. Y11282) were amplified by PCR.

Figure 25 shows the results of amplification using the SBE II-Intron 5 primer set (primer set 6: sr913F and WBE2E6 R) on various diploid, tetraploid and hexaploid 25 wheats.

- i) *T.boeodicum* (A genome diploid)
- ii) *T.tauschii* (D genome diploid)
- iii) *T.aestivum* cv. Chinese Spring ditelosomic line 2AS (lacking chromosome arm 2AL)
- 30 iv) Crete 10 (AABB tetraploid)
- v) *T. aestivum* cv Rosella (hexaploid)

The horizontal axis indicates the size of the product in base pairs, the vertical axis shows arbitrary fluorescence units. The various arrows indicate the products of different genomes: A, A genome, B, B genome, D, D genome, 35 U, unassigned additional product.

- 18 -

Figure 26 shows the results obtained by amplification using the SBE II-Intron 10 primer set (primer set 11: da5.seq and WBE2E11R on the wheat lines:

5 (i) *T. aestivum* cv. Chinese Spring ditelosomic line 2AS.

(ii) *T. aestivum* Chinese Spring nullisomic/tetrasomic line N2BT2A.

(iii) *T. aestivum* Chinese Spring nullisomic/tetrasomic line N2DT2B.

10 The horizontal axis indicates the size of the product in base pairs, the vertical axis shows arbitrary fluorescence units. The various arrows indicate the products of different genomes: A, A genome, B, B genome, D, D genome.

Figure 27 shows the results of transient  
15 expression assays typical of each promoter and target tissue. The photographs (40 x magnification) of representative tissue resulting from the transient expression assays typical of each promoter and target tissue revealed under a Leica microscope with blue light  
20 illumination. Photographs were taken 48 to 72 hours after tissue bombardment. The promoter constructs are listed as follows, (with the panels showing endosperm, embryo and leaf expression listed in respective order): pact\_jsgfp\_nos (panels a, g and m); pwsssIpro1gfpNOT (panels b, h and n);  
25 pwsssIpro2gfpNOT (panels c, i and o); psbeIIpro1gfpNOT (panels d, j and p); psbeIIpro2gfpNOT (panels e, k and q); pZLgfpNOT (Panels f, l and r).

#### Example 1 Identification of Gene Encoding SBE I

#### 30 **Construction of Genomic Library and Isolation of Clones**

The genomic library used in this study was constructed from *Triticum tauschii*, var. *strangulata*, accession number CPI 100799. Of all the accessions of *T. tauschii* surveyed, the genome of CPI 100799 is the most  
35 closely related to the D genome of hexaploid wheat.



*Triticum tauschii*, var *strangulata* (CPI accession number 110799) was kindly provided by Dr E Lagudah. Leaves were isolated from plants grown in the glasshouse.

DNA was extracted from leaves of *Triticum tauschii* using published methods (Lagudah et al, 1991), partially digested with *Sau3A*, size fractionated and ligated to the arms of lambda GEM 12 (Promega). The ligated products were used to transfect the methylation-tolerant strain PMC 103 (Doherty et al. 1992). A total of  $2 \times 10^6$  primary plaques were obtained with an average insert size of about 15 kb. Thus the library contains approximately 6 genomes worth of *T. tauschii* DNA. The library was amplified and stored at 4°C until required.

Positive plaques in the genomic library were selected as those hybridising with the 5' end of a maize starch branching enzyme I cDNA (Baba et al, 1991) using moderately stringent conditions as described in Rahman et al, (1997).

#### 20 **Preparation of Total RNA from Wheat**

Total RNA was isolated from leaves, pre-anthesis pericarp and different developmental stages of wheat endosperm of the cultivar, Hartog and Rosella. This material was collected from both the glasshouse and the field. The method used for RNA isolation was essentially the same as that described by Higgins et al (1976). RNA was then quantified by UV absorption and by separation in 1.4% agarose-formaldehyde gels which were then visualized under UV light after staining with ethidium bromide (Sambrook et al, 1989).

#### **DNA and RNA analysis**

DNA was isolated and analysed using established protocols (Sambrook et al, 1989). DNA was extracted from wheat (cv. Chinese Spring) using published methods (Lagudah et al, 1991). Southern analysis was performed essentially as described by Jolly et al (1996). Briefly, 20 µg wheat

DNA was digested, electrophoresed and transferred to a nylon membrane. Hybridisation was conducted at 42°C in 25% or 50% formamide, 2 x SSC, 6% Dextran Sulphate for 16h and the membrane was washed at 60°C in 2 x SSC for 3 x 1h unless  
5 otherwise indicated. Hybridisation was detected by autoradiography using Fuji X-Omat film.

RNA analysis was performed as follows. 10 µg of total RNA was separated in a 1.4% agarose-formaldehyde gel and transferred to a nylon Hybond N<sup>+</sup> membrane (Sambrook et  
10 al, 1989 ), and hybridized with cDNA probe at 42°C in Khandjian hybridizing buffer (Khandjian, 1989). The 3' part of wheat SBE I cDNA (designated wSBE I-D43, see Table 1) was labelled with the Rapid Multiprime DNA Probe Labelling Kit (Amersham) and used as probe. After washing at 60°C with  
15 2 x SSC, 0.1% SDS three times, each time for about 1 to 2 hours, the membrane was visualized by overnight exposure at -80°C with X-ray film, Kodak MR.

#### Example 2

#### Frequency of Recovery of SBE I Type Clones from the Genomic Library

20 An estimated  $2 \times 10^6$  plaques from the amplified library were screened using an *EcoRI* fragment that contained 1200 bp at the 5' end of maize SBE I (Baba et al, 1991) and twelve independent isolates were recovered and purified.  
25 This corresponds to the screening of somewhat fewer than the  $2 \times 10^6$  primary plaques that exist in the original library (each of which has an average insert size of 15 kb) (Maniatis et al, 1982), because the amplification may lead to the representation of some sequences more than others.  
30 Assuming that the amplified library contains approximately three genomes of *T. tauschii*, the frequency with which SBE I-positive clones were recovered suggests the existence of about 5 copies of SBE I type genes within the *T. tauschii* genome.

35 Digestion of DNA from the twelve independent isolates by the restriction endonuclease *BamHI* followed by hybridisation with a maize SBE I clone, suggested that the

- 21 -

genomic clones could be separated into two broad classes (Figure 1). One class had 10 members and a representative from this class is the clone  $\lambda$ E1 (Figure 1, lane 1);  $\lambda$ E6 (Figure 1, lane 3) is a member of this class, but is missing the 5' end of the E1-SBE I gene because the SBE I gene is at the extremity of the cloned DNA. Further hybridisation studies at high stringency with the extreme 5' and 3' regions of the SBE I gene contained in  $\lambda$ E1 suggested that the other clones contained either identical or very closely related genes.

The second family had two members, and of these clone  $\lambda$ E7 (Figure 1, lane 4) was arbitrarily selected for further study. These two members did not hybridise to probes from the extreme 5' and 3' regions of the SBE I gene that were contained in  $\lambda$ E1, indicating that they were a distinct sub-class.

The DNA from *T. tauschii* and the lambda clones  $\lambda$ E1 and  $\lambda$ E7 was digested with *Bam*HI and hybridised with fragment E1.1, as shown in Figure 2. This fragment contains sequences that are highly conserved (85% sequence identity over 0.3 kb between  $\lambda$ E1 and  $\lambda$ E7), corresponding to exons 3, 4 and 5 of the rice gene. The bands in the genomic DNA at 0.8 kb and 1.0 kb correspond to identical sized fragments from  $\lambda$ E1 and  $\lambda$ E7, as shown in Figure 2; these are fragments E1.1 and E7.8 of  $\lambda$ E1 and  $\lambda$ E7 genomic clones respectively. Thus the arrangement of genes in the genomic clones is unlikely to be an artefact of the cloning procedure. There are also bands in the genomic DNA of approximately 2.5 kb, 4.8 kb and 8 kb in size which are not found from the digestion of  $\lambda$ E1 or  $\lambda$ E7; these could represent genes such as the 5' sequences of wSBE I-D1 or wSBE I-D3; see below.

Example 3                      Tandem Arrangement of SBE I Type Genes in the *T. tauschii* Genome

Basic restriction endonuclease maps for  $\lambda$ E1 and  $\lambda$ E7 are shown in Figure 3. The map was constructed by

- 22 -

performing a series of hybridisations of *EcoRI* or *BamHI* digested DNA from  $\lambda$ E1 or  $\lambda$ E7. The probes used were the fragments generated from *BamHI* digestion of the relevant clone. Confirmation of the maps was obtained by PCR  
5 analysis, using primers both within the insert and also from the arms of lambda itself. PCR was performed in 10  $\mu$ l volume using reagents supplied by Perkin-Elmer. The primers were used at a concentration of 20  $\mu$ M. The program used was 94°C, 2 min, 1 cycle, then 94°C, 30 sec; 55°C, 30 sec; 72°C,  
10 1min for 36 cycles and then 72°C, 5 min; 25°C, 1 min.

Sequencing was performed on an ABI sequencer using the manufacturer's recommended protocols for both dye primer and dye terminator technologies. Deletions were carried out using the Erase-a-base kit from Promega.

15 Sequence analysis was carried out using the GCG version 7 package of computer programs (Devereaux et al, 1984).

The PCR products were also used as hybridisation probes. The positioning of the genes was derived from  
20 sequencing the ends of the *BamHI* subclones and also from sequencing PCR products generated from primers based on the insert and the lambda arms. The results indicate that there is only a single copy of a SBE I type gene within  $\lambda$ E1. However, it is clear that  $\lambda$ E7 resulted from the cloning of a  
25 DNA fragment from within a tandem array of the SBE I type genes. Of the three genes in the clone, which are named as WSBE I-D1, WSBE I-D2 and WSBE I-D3); only the central one (WSBE I-D2) is complete.

#### 30 Example 4      Construction and Screening of cDNA Library

A wheat cDNA library was constructed from the cultivar Rosella using pooled RNA from endosperm at 8, 12, 18 and 20 days after anthesis.

35 The cDNA library was prepared from poly A<sup>+</sup> RNA that was extracted from developing wheat grains (cv. Rosella, a hexaploid soft wheat cultivar) at 8, 12, 15, 18, 21 and 30 days after anthesis. The RNA was pooled and used

- 23 -

to synthesise cDNA that was propagated in lambda ZapII (Stratagene).

The library was screened with a genomic fragment from  $\lambda$ E7 encompassing exons 3, 4 and 5 (fragment E7.8 in Figure 3). A number of clones were isolated. Of these an apparently full-length clone appeared to encode an unusual type of cDNA for SBE I. This cDNA has been termed SBE I-D2 type cDNA. The putative protein product is compared with the maize SBE I and rice SBE I type deduced amino acid sequences in Figure 4. The main difference is that this putative protein product is shorter at the C-terminal end, with an estimated molecular size of approximately 74 kD compared with 85 kDa for rice SBE I (Kawasaki *et al*, 1993). Note that amino acids corresponding to exon 9 of rice are missing in SBE I-D2 type cDNA, but those corresponding to exon 10 are present. There are no amino acid residues corresponding to exons 11-14 of rice; furthermore, the sequence corresponding to the last 57 amino acids of SBE I-D2 type has no significant homology to the sequence of the rice gene.

We expressed SBE I-D2 type cDNA in *E. coli* in order to examine its function. The cDNA was expressed as a fusion protein with 22 N-terminal residues of  $\beta$ -galactosidase and two threonine residues followed by the SBE I-D2 cDNA sequence either in or out of frame. Although an expected product of about 75 kDa in size was produced from only the in-frame fusion, we could not detect any enzyme activity from crude extracts of *E. coli* protein. Furthermore the in-frame construct could not complement an *E. coli* strain with a defined deletion in glycogen branching, although other putative branching enzyme cDNAs have been shown to be functional by this assay (data not shown). It is therefore unclear whether the wSBE I-D2 gene in  $\lambda$ E7 codes for an active enzyme *in vivo*.

Example 5                      Gene Structure in E7

**i. Sequence of wSBE I-D2**

We sequenced 9.2 kb of DNA that contained wSBE I-D2. This corresponds to fragments 7.31, 7.8 and 7.18. Fragment 7.31 was sequenced in its entirety (4.1 kb), but the sequence of about 30 bases about 2 kb upstream of the start of the gene could not be obtained because it was composed entirely of Gs. Elevation of the temperature of sequencing did not overcome this problem. Fragments 7.8 (1 kb) and 7.18 (4 kb) were completely sequenced, and corresponded to 2 kb downstream of the last exon detected for this gene. It was clear that we had isolated a gene which was closely related (approximately 95% sequence identity) to the SBE I-D2 type cDNA referred to above, except that the last 200 bp at the 3' end of the cDNA are not present. The wSBE I-D2 gene includes sequences corresponding to rice exon 11 which are not in the cDNA clone. In addition it does not have exons 9, 12, 13 or 14; these are also absent from the SBE I-D2 type cDNA. The first two exons show lower identity to the corresponding exons from rice (approximately 60%) (Kawasaki et al, 1993) than to the other exons (about 80%). A diagrammatic exon-intron structure of the wSBE I-D2 gene is indicated in Figure 5. The restriction map was confirmed by sequencing the PCR products that spanned fragments 7.18 and 7.8 and 7.8 and E7.31 (see Figure 3) respectively.

**ii. Sequence of wSBE I-D3**

This gene was not sequenced in detail, as the genomic clone did not extend far enough to include the 5' end of the sequence. The sequence is of a SBE-I type. The orientation of the gene is evident from sequencing of the relevant *Bam*HI fragments, and was confirmed by sequence analysis of a PCR product generated using primers from the right arm of lambda and a primer from the middle of the gene. The sequence homology with wSBEI-D2 is about 80% over the regions examined. The 2 kb sequenced corresponded to

exons 5 and 6 of the rice gene; these sequences were obtained by sequencing the ends of fragments 7.5, 7.4 and 7.14 respectively, although the sequences from the left end of fragment 7.14 did not show any homology to the rice sequences. The gene does not appear to share the 3' end of SBE I-D2 type cDNA, as a probe from 500 bp at the 3' end of the cDNA (including sequences corresponding to exons 8 and 10 from rice) did not hybridise to fragment 7.14, although it hybridised to fragment 7.18.

### iii. Sequence of wSBE I-D1

This gene was also not sequenced in detail, as it was clear that the genomic clone did not extend far enough to include the 5' sequences. Limited sequencing suggests that it is also a SBE I type gene. The orientation relative to the left arm of lambda was confirmed by sequencing a PCR product that used a primer from the left arm of lambda and one from the middle of the gene (as above). Its sequence homology with wSBE I-D2 ,D3 and D4 (see below) is about 75% in the region sequenced corresponding to a part of exon 4 of the rice gene.

Starch branching enzymes are members of the  $\alpha$ -amylase protein family, and in a recent survey Svensson (1994) identified eight residues in this family that are invariant, seven in the catalytic site and a glycine in a short turn. Of the seven catalytic residues, four are changed in SBE I-D2 type. However, additional variation in the 'conserved' residues may come to light when more plant cDNAs for branching enzyme I are available for analysis. In addition, although exons 9, 11, 12, 13 and 14 from rice are not present in the SBE I-D2 type cDNA, comparison of the maize and rice SBE I sequences indicate that the 3' region (from amino acid residue 730 of maize) is much more variable than the 5' and central regions. The active sites of rice and maize SBE I sequences, as indicated by Svensson (1994), are encoded by sequences that are in the central portion of the gene. When SBE II sequences from *Arabidopsis* were

- 26 -

compared by Fisher *et al* (1996) they also found variation at the 3' and 5' ends. SBE I-D2 type cDNA may encode a novel type of branching enzyme whose activity is not adequately detected in the current assays for detecting branching enzyme activity; alternatively the cDNA may correspond to an endosperm mRNA that does not produce a functional protein.

Example 6                      Cloning of the cDNA corresponding to the wSBE I-D4 gene

The first strand cDNAs were synthesized from 1 µg of total RNA, derived from endosperm 12 days after pollination, as described by Sambrook *et al* (1989), and then used as templates to amplify two specific cDNA regions of wheat SBE I by PCR.

Two pairs of primers were used to obtain the cDNA clones BED1 and BED3 (Table 1). Primers used for cloning of BED3 were the degenerate primer NTS5'

5' GGC NAC NGC NGA G/AGA C/TGG 3' (SEQ ID NO.1),

based on the N-terminal sequence of the purified wheat endosperm SBE I protein, in which the 5' end of the primer is at position 168 of wSBE I-D4 cDNA, as shown in Table 1, based on the N-terminal sequence of wheat SBE I, and the primer NTS3'.

5' TAC ATT TCC TTG TCC ATCA 3' (SEQ ID NO.2)

in which the 5' end is at position 1590 of wSBE I-D4 cDNA, (see Table 1), designed to anneal to the conserved regions of the nucleotide sequences of BED5 and the maize and rice SBE I cDNAs. For clone BED1, the primers used were BEC5'

5' ATC ACG AGA GCT TGC TCA (SEQ ID NO.3)



- 27 -

in which the 5' end is at position 1 of wSBE I-D4 cDNA (see Table 1); the sequence was based on the wSBE I-D4 gene, and BEC3'

5 5' CGG TAC ACA GTT GCG TCA TTT TC 3' (SEQ ID NO.4)

in which the 5' end is at position 334 of wSBE I-D4 cDNA (see Table 1), and the sequence was based on BED 3.

10

Example 7                      Identification of the gene from the *Triticum tauschii* SBE I family which is expressed in the endosperm

We have isolated two classes of SBE I genomic clones from *T. tauschii*. One class contained two genomic clone isolates, and this class has been characterised in some detail (Rahman et al, 1997). The complete gene contained within this class of clones was termed wSBE I-D2; there were additional genes at either ends of the clone, and these were designated wSBE I-D1 and wSBE I-D3. The other class contained nine genomic clone isolates. Of these λE1 was arbitrarily taken as a representative clone, and its restriction map is shown in Figure 3; the SBE I gene contained in this clone was called wSBE I-D4.

Fragments E1.1 (0.8 kb) and E1.2 (2.1 kb) and fragments E1.7 (4.8 kb) and E1.5 (3 kb) respectively were completely sequenced. Fragment E1.7 was found to encode the N-terminal of the SBE I, which is found in the endosperm as described in Morell et al (1997). This is shown in Figure 6. Using antibodies raised against the N-terminal sequence, Morell et al (1997) found that the D genome isoform was the most highly expressed in the cultivars Rosella and Chinese Spring. We have thus isolated from *T. tauschii* a gene, wSBE I-D4, whose homologue in the hexaploid wheat genome encodes the major isoform for SBE I that is found in the wheat endosperm.

- 28 -

**Table 1**

**Location of structural features and probes within wSBE I-D4 sequence.**

- 5 A. Location of exons by comparison with the cDNA sequence of Repellin et al., (1997). Accession number Y12320.

	Exon number	Start posn	End posn
10	1	4890	4987
	2	5082	5149
	3	5524	5731
	4	5819	5888
	5	6149	6318
15	6	6519	7424
	7	7744	7860
	8	8015	8077
	9	8562	8670
20	10	9137	9237
	11	9421	9488
	12	9580	9661
	13	9781	9897
	14	9990	10480

- 25 B. Other features.

	Name of feature.	wSBE I-D4. sequence	D4 cDNA sequence.
30	Putative initiation of translation	4900	11
	Mature N-terminal sequence of SBE I	5550	124
	End of translated SBE I sequence	10225	2431
	End of D4 cDNA sequence	10461	2687
	wSBE I-D45	4870, 5860	1, 354
35	wSBE I-D43	10116, 10435	2338, 2657
	E1.1	5680, 6400	380, 630
	BED 1		1, 354
	BED 2		169, 418
	BED 3		151, 1601
40	BED 4		867, 2372
	BED 5		867, 2687
	Endosperm box like motif TGAAAAGT	4480, 590	
	CAAAT motif	4863	
	TATAAA motif	4833	

- 29 -

All nine genomic clones of the  $\lambda$ E1 type isolated from *T. tauschii* appear to contain the *wSBE I-D4* gene, or very similar genes, on the basis of PCR amplification and hybridisation experiments. However, the restriction patterns obtained for the clones differ with *Bam*HI and *Eco*RI, among other enzymes, indicating that either the clones represent near-identical but distinct genes or they represent the same gene isolated in distinct products of the *Sau*3A digest used to generate the library.

Example 8                      Investigation of other SBE I genomic clones isolated

All ten members of the  $\lambda$ E1-like class of SBE I genomic clones were investigated by hybridisation with probes derived from fragment E1.7 (sequence *wSBE I-D45*, encoding the translation start signal and the first 100 amino acids from the N-terminal end and intron sequences; see Table 1) and from fragment E1.5 (sequence *wSBE I-D43*, corresponding largely to the 3' untranslated sequence and containing intron sequences, see Table 1). The results obtained were consistent with one type of gene being isolated in different fragments in the different clones, as shown in Figure 7. The PCR products were obtained from the clones  $\lambda$ E1, 2, 9, 14, 27, 31 and 52. These hybridised to *wSBE I-D45* using primers that amplify near the 5' end of the gene (positions 5590-6162 of *wSBE I-D4*). Sequencing showed no differences in sequence of a 200 bp product.

Analysis of the promoter for *wSBE I-D4* allows us to investigate the presence of motifs previously described for promoters that regulate gene expression in the endosperm. Forde et al (1985) compared prolamin promoters, and suggested that the presence of a motif approximately -300 bp upstream of the transcription start point, called the endosperm box, was responsible for endosperm-specific expression. The endosperm box was subsequently considered to consist of two different motifs: the endosperm motif (EM) (canonical sequence TGTAAG) and the GCN 4 motif (canonical

- 30 -

sequence G/ATGAG/CTCAT). The GCN4 box is considered to regulate expression according to nitrogen availability (Muller and Knudsen, 1993). The *wSBE I-D4* promoter contains a number of imperfect EM-like motifs at approximately -100, -300 and -400 as well as further upstream. However, no GCN4 motifs could be found, which lends support to the idea that this motif regulates response to nitrogen, as starch biosynthesis is not as directly dependent on the nitrogen status of the plant as storage protein synthesis. Comparison of the promoters for *wSBE I-D4* and *D2* (Rahman et al, 1997) indicates that although there are no extensive sequence homologies there is a region of about 100 bp immediately before the first encoded methionine where the homology is 61% between the two promoters. In particular there is an almost perfect match in the sequence over twenty base pairs CTCGTTGCTTCC/TACTCCACT, (positions 4723-4742 of the *wSBE I* sequence), but the significance of this is hard to gauge, as it does not occur in the rice promoter for *SBE I*. The availability of more promoters for starch biosynthetic enzymes may allow firmer conclusions to be drawn. There are putative CAAT and TATA motifs at positions 4870 and 4830 respectively of *wSBE I-D4* sequence. The putative start of translation of the mRNA is at position 4900 of *wSBE I-D4*.

Figure 5 shows the structure of the *wSBE I-D4* gene, compared with the genes from rice and wheat (Kawasaki et al, 1993; Rahman et al, 1997). The rice *SBE I* has 14 exons compared with 13 for *wSBE I-D4* and 10 for *wSBE I-D2*. There is good conservation of exon-intron structure between the three genes, except at the extreme 5' end. In particular the sizes of intron 1 and intron 2 are very different between rice *SBE I* and *wSBE I-D4*.

#### Example 9                      Isolation of cDNA for SBE I

Using the maize starch branching enzyme I cDNA as a probe (Baba et al, 1991), 10 positive plaques were recovered by screening approximately  $10^5$  plaques from a wheat endosperm cDNA library prepared from the cultivar

- 31 -

Rosella, as described in Example 4. On purifying and sequencing these plaques it was clear that even the longest clone (BED5, 1822 bp) did not encode the N-terminal sequence obtained from protein analysis. Degenerate primers based on the wheat endosperm SBE I protein N-terminal sequence (Morell et al, 1997) and the sequence from BED5 were then used to amplify the 5' region: this produced a cDNA clone termed BED 3 (Table 1 and Figure 8). This cDNA clone overlapped extensively and had 100% sequence identity with BED5 and BED4 (Figure 8). As almost the entire protein N-terminal sequence had been included in the primer sequence design, this did not provide independent evidence of the selection of a cDNA sequence in the endosperm that encoded the protein sequence of the main form of SBE I. Using a BED3 to screen a second cDNA library produced BED2, which is shorter than BED3 but confirmed the BED3 sequence at 100% identity between positions 169 and 418 (Figure 8 and Table 1). In addition the entire cDNA sequence for BED3 could be detected at a 100% match in the genomic clone  $\lambda$ E1. Primers based on the putative transcription start point combined with a primer based on the incomplete cDNAs recovered were then used to obtain a PCR product from total endosperm RNA by reverse transcription. This led to the isolation of the cDNA clone, BED1, of 300 bp, whose location is shown in Figure 8. By analysing this product, a sequence was again obtained that could be found exactly in the genomic clone  $\lambda$ E1, and which overlapped precisely with BED3.

The N-terminal of the protein matches that of SBE I isolated from wheat endosperm by Morell et al (1997), and thus the *wSBE I-D4* cDNA represents the gene for the predominant SBE I isoform expressed in the endosperm. The encoded protein is 87 kDa; this is similar to proteins encoded by maize (Baba et al, 1991) and rice (Nakamura et al, 1992) cDNAs for SBE I and is distinct from the *wSBE I-D2* cDNA described previously, in which the encoded protein was 74 kDa (Rahman et al, 1997).

- 32 -

Five cDNA clones were sequenced and their sequences were assembled into one contiguous sequence using a GCG program (Devereaux et al, 1984). The arrangement of these sequences is illustrated in Figure 8, the nucleotide sequence is shown in SEQ ID No:5, and the deduced amino acid sequence is shown in SEQ ID No:6. The intact cDNA sequence, *wsBE I-D4* cDNA, is 2687 bp and contains one large open reading frame (ORF), which starts at nucleotides 11 to 13 and ends at nucleotides 2432 to 2434. It encodes a polypeptide of 807 amino acids with a molecular weight of 87 kDa. Comparison of the amino acid sequence encoded by *wsBE I-D4* cDNA with that encoded by maize and rice *SBE I* cDNAs showed that there is 75-80% identity between any of two these sequences at the nucleotide level and almost 90% at the amino acid level. Alignment of these three polypeptide sequences, as shown in Figure 4, along with the deduced sequences for pea, potato and *wsBE I-D2* type cDNA, indicated that the sequences in the central region are highly conserved, and sequences at the 5' end (about 80 amino acids) and the 3' end (about 60 amino acids) are variable.

Svensson et al (1994) indicated that there were several invariant residues in sequences of the  $\alpha$ -amylase super-family of proteins to which *SBE I* belongs. In the sequence of maize *SBE I* these are in motifs commencing at amino acid residue positions 341, 415, 472, 537 respectively; these are also encoded in the *wsBE I-D4* sequence (SEQ ID No:9), further supporting the view that this gene encodes a functional enzyme. This is in contrast to the results with the *wsBE I-D2* gene, where three of the conserved motifs appear not to be encoded (Rahman et al, 1997).

There is about 90% sequence identity in the deduced amino acid sequence between *wsBE I-D4* cDNA and rice *SBE I* cDNA in the central portion of the molecule (between residues 160 and 740 for the deduced amino acid product from *wsBE I-D4* cDNA). The sequence identity of the deduced amino

- 33 -

acid sequence of the *wsBE I-D4* cDNA to the deduced amino acid sequence of *wsBE I-D2* is somewhat lower (85% for the most conserved region, between residues 285 to 390 for the deduced product of *wsBE I-D4* cDNA). Surprisingly, however, *wsBE I-D4* cDNA is missing the sequence that encodes amino acids at positions 30 to 58 in rice SBE I (see Figure 4). This corresponds to residues within the transit peptide of rice SBE I. A corresponding sequence also occurs in the deduced amino acid sequence from maize *SBE I* (Baba et al, 1991) and *wsBE I-D2* type cDNA (Rahman et al, 1997). Consequently the transit sequence encoded by *wsBE I-D4* cDNA is unusually short, containing only 38 amino acids, compared with 55-60 amino acids deduced for most starch biosynthetic enzymes in cereals (see for example Ainsworth, 1993; Nair et al, 1997). The *wsBE I-D4* gene does contain this sequence, but this does not appear to be transcribed into the major species of RNA from this gene, although it can be detected at low relative abundance. This raises the possibility of alternative splicing of the *wsBE I-D4* transcript, and also the question of the relative efficiency of translation/transport of the two isoforms. The possibility of alternative splicing in both rice and wheat has been considered for soluble starch synthase (Baba et al, 1993; Rahman et al, 1995). Alternative splicing of soluble starch synthase would give a transit sequence of 40 amino acids, which is the same length proposed for the product of *wsBE I-D4* cDNA.

We have previously used probes based on exons 4, 5 and 6 (E7.8 and E1.1, see Rahman et al., 1997) of *wsBE-D2* to probe wheat and *T. tauschii* genomic DNA cleaved with *PvuII* and *BamHI* respectively. This region is highly conserved within rice *SBE I*, *wsBE I-D2* and *wsBE I-D4* and produced ten bands with wheat DNA and five with *T. tauschii* DNA. Neither *PvuII* nor *BamHI* cleaved within the probe sequences, suggesting that each band represented a single type of SBE I gene. We have described four SBE I genes from *T. tauschii*: *wsBE I-D1*, *wsBE I-D2*, *wsBE I-D3* and *wsBE I-D4* (Rahman et al,

- 34 -

1997 and this specification), and so we may have accounted for most of the genes in *T. tauschii* and, by extension, the genes from the D genome of wheat. In wheat, at least two hybridising bands could be assigned to each of  
5 chromosomes 7A, 7B and 7D.

Example 10                      Tissue specificity and expression during endosperm development

The 300 bp of 3' untranslated sequence of  
10 *wsBE I-D4* cDNA does not show any homology with either the *wsBE I-D2* type cDNA that we have described earlier (Rahman et al, 1997) or with BE-I from rice, as shown in Figure 5. We have called this sequence *wsBE I-D43C* (see SEQ ID No:9). It seemed likely that *wsBE I-D43C* would be a specific probe  
15 for this class of SBE-I, and thus it was used to investigate the tissue specificity. Hybridization of RNA from endosperm of hexaploid *T. tauschii* cultures with SBE I, SBE II, SSS I, DBE I, wheat actin, and wheat ribosomal RNA was examined. RNA was purified at various numbers of days after anthesis  
20 from plants grown with a 16 h photoperiod at 13 °C (night) and 18 °C (day). The age of the endosperms from which RNA was extracted in days after anthesis is given above the lanes in the blot. Equivalent amounts of RNA were electrophoresed in each lane. The probes used are identified  
25 in Tables 1 and 2.

The results are shown in Figures 9a to 9g. An RNA species of about 2700 bases in size was found to hybridise. This is very close to the size of the *wsBE I-D4* cDNA sequence. RNA hybridising to *wsBE-I-D43C* is most abundant  
30 at the mid-stage of endosperm development, as shown in Figure 9a, and in field grown material is relatively constant during the period 12-18 days, the time at which there is rapid starch and storage protein accumulation (Morell et al, 1995).

35 The sequence contained within the *wsBE I-D4* gene appears to be expressed only in the endosperm (Figure 9a, Figure 9b). We could not detect any expression in the leaf.



- 35 -

This could be because another isoform is expressed in the leaf, and/or because the amount of SBE I present in the leaf is much less than what is required in the endosperm. Isolation of SBE I clones from a leaf cDNA library would  
5 enable this question to be resolved.

Example 11            Intron-Exon Structure of SBE I

By comparison of the cDNA sequence of SBE I (Repellin et al, 1997) with that of *wSBE I-D4* we can deduce  
10 the intron-exon structure of the gene for the major isoform of SBE I that is found in the endosperm. The structure contains 14 exons compared to 14 for rice (Kawasaki et al, 1993). These 14 exons are spread over 6 kb of sequence, a distance similar to that found in both rice *SBE I* and  
15 *wSBE I-D2*. A dotplot comparison of *wSBE I-D4* sequence and that of rice *SBE I* sequence, depicted in Figure 10, shows good sequence identity over almost the entire gene starting from about position 5100 of *wSBE I-D4*; the identity is poor over the first 5 kb of sequence corresponding largely to the  
20 promoter sequences. The sequence identity over introns (about 60%) is lower than over exons (about 85%).

Example 12            Repeated Sequences in SBE I

Sequencing of *wSBE I-D4* revealed there was a  
25 repeated sequence of at least 300 bp contained in a 2kb fragment about 600 bp after the 3' end of the gene. We have called this sequence *wSBE I-D4R* (SEQ ID NO: 9). This repeated sequence is within fragment E1.5 (Figure 3 and Table 1) and is flanked by non-repetitive sequences from the  
30 genomic clone. We have previously shown that the restriction pattern obtained by digesting  $\lambda$ E1 with the restriction enzyme *Bam*HI is also obtained when *T. tauschii* DNA is digested. Thus *wSBE I-D4R* is unlikely to be a cloning artefact. A search of the GenBank Database revealed  
35 that *wSBE I-D4R* shared no significant homology with any sequence in the database. Hybridisation experiments with *wSBE I-D4R* showed that all of the other *SBE I-D4* type

- 36 -

genomic clones (except number 29) contained this repeated sequence (data not shown). The *wSBE I-D4R* sequence was not highly repeated and occurred in the wheat genome with a similar frequency as the *wSBE I-D4* sequence.

5           When *SBE I-D4R* was used as the probe on wheat DNA from the nulli-tetra lines, four bands were obtained; two of these bands could be assigned to chromosome 7A and the others to chromosomes 7B and 7D (Figure 11). One of the two *Bam*HI fragments from wheat DNA which could be assigned to  
10 chromosome 7A was distinct from the single band from chromosome 7A detected using *wSBE I-D43* as the probe; the other three bands coincided in the autoradiograph with bands obtained with *wSBE I-D43*, and are likely to represent the same fragment. However, one of these fragments was distinct  
15 from the *Bam*HI fragment that hybridised to the *wSBE I-D43* sequence. In *wSBE I-D4* (see SEQ ID No:9), the *wSBE I-D43* sequence is only 300 bp upstream of *wSBE I-D4R*, and occurs in the same *Bam*HI fragment. These results suggest that the *wSBE I-D4R* sequence can occur independently of *wSBE I-D4* in  
20 the wheat genome.

#### Example 13           Isolation of Genomic Clones Encoding SBE II

Screening of a cDNA library, prepared from the wheat endosperm as described in Example 4, with the maize  
25 BE I clone (Baba et al, 1991) at low stringency led to the isolation of two classes of positive plaques. One class was strongly hybridising, and led to the isolation of wheat *SBE I-D2* type and *SBE I-D4* type cDNA clones, as described in Example 5 and in Rahman et al (1997). The second class was  
30 weakly hybridising, and one member of this class was purified. This weakly hybridising clone was termed *SBE-9*, and on sequencing was found to contain a sequence that was distinct from that for *SBE I*. This sequence showed greatest homology to maize BE II sequences, and was considered to  
35 encode part of the wheat *SBE II* sequence.

The screening of approximately  $5 \times 10^5$  plaques from a genomic library constructed from *T. tauschii* (see

- 37 -

Example 1) with the SBE-9 sequence led to the isolation of four plaques that were positive. These were designated *wSBE II-D1* to *wSBE II-D4* respectively, and were purified and analysed by restriction mapping. Although they all had different hybridization patterns with SBE-9, as shown in Figure 12, the results were consistent with the isolation of the same gene in different-sized fragments.

Example 14                      Identification of the N-terminal sequence of SBE II

Sequencing of the SBE II gene contained in clone 2, termed *SBE II-D1* (see SEQ ID No:10), showed that it coded for the N-terminal sequence of the major isoform of SBE II expressed in the wheat endosperm, as identified by Morell et al (1997). This is shown in Figure 13.

Example 15                      Intron-Exon Structure of the SBE II Gene

In addition to encoding the N-terminal sequence of SBE II, as shown in Example 10, the cDNA sequence reported by Nair et al (1997) was also found to have 100% sequence identity with part of the sequence of *wSBE II-D1*. Thus the intron-exon structure can be deduced, and this is shown in Figure 14. The positions of exons and other major structural features of the SBE II gene are summarized in Table 2.

Example 16                      Number of SBE II Genes in *T. tauschii* and Wheat

Hybridisation of the SBE II conserved region with *T. tauschii* DNA revealed the presence of three gene classes. However, in our screening we only recovered one class. Hybridisation to wheat DNA indicated that the locus for SBE II was on chromosome 2, with approximately 5 loci in wheat; most of these appear to be on chromosome 2D, as shown in Figure 15.

- 38 -

**Table 2**  
**Positions of structural features in wSBE II-D1.**

5 A. Positions of exons.

	Exon number	Genomic start	Genomic finish
10	1	1058	1336
	2	1664	1761
	3	2038	2279
	4	2681	2779
	5	2949	2997
15	6	3145	3204
	7	3540	3620
	8	3704	3825
	9	4110	4188
	10	4818	4939
20	11	5115	5234
	12	6209	6338
	13	6427	6549
	14	6739	6867
	15	7447	7550
25	16	8392	8536
	17	9556	9703
	18	9839	9943
	19	10120	10193
	20	10395	10550
30	21	10928	11002
	22	11092	11475

B. Other structural features within the wSBE II-D1 DNA  
sequence

35	Putative initiation of translation	1214
	Mature N-terminal sequence of SBE II.	1681
	wSBE II-D13	11116 to 11448
40	Endosperm box like motif TGAAAAGT	521
	Endosperm box like motif TGAAAGT	565
	Endosperm box like motif CGAAAAT	669
	Endosperm box like motif TAAATGT	768
	CAAAAT motif	784
45	TCAATT motif	1108
	TATAAA motif	799
	AATTAA motif	1110

### Example 17      Expression of SBE II

Investigation of the pattern of expression of SBE II revealed that the gene was only expressed in the endosperm. However the timing of expression was quite  
5 distinct from that of SBE I, as illustrated in Figures 9a, 9b and 9c.

SBE I gene expression is only clearly detectable from the mid-stage of endosperm development (10 days after anthesis in Figure 9b), whereas SBE II gene expression is  
10 clearly seen much earlier, in endosperm tissue at 5-8 days after development (Figures 9a and 9c), corresponding to an early stage of endosperm development. The hybridisation of wheat endosperm mRNA with the actin and ribosomal RNA genes is shown as controls (Figures 9fa and 9g, respectively).

15

### Example 18      Cloning of Wheat Soluble Starch Synthase cDNA

A conserved sequence region was used for the synthesis of primers for amplification of SSS I by  
20 comparison with the nucleotide sequences encoding soluble starch synthases of rice and pea. A 300 bp RT-PCR product was obtained by amplification of cDNA from wheat endosperm at 12 days post anthesis. The 300 bp RT-PCT product was then cloned, and its sequence analysed. The comparison of  
25 its sequence with rice SSS cDNA showed about 80% sequence homology. The 300 bp RT-PCR product was 100% homologous to the partial sequence of a wheat SSS I in the database produced by Block et al (1997).

The 300 bp cDNA fragment of wheat soluble starch  
30 synthase thus isolated was used as a probe for the screening of a wheat endosperm cDNA library (Rahman et al, 1997). Eight cDNA clones were selected. One of the largest cDNA clones (sm2) was used for DNA sequencing analysis, and gave a 2662 bp nucleotide sequence, which is shown in SEQ ID  
35 NO:14. A large open reading frame of this cDNA encoded a 647 amino acid polypeptide, starting at nucleotides 247 to 250 and terminating at nucleotides 2198 to 2200. The

- 40 -

deduced polypeptide was shown by protein sequence analysis to contain the N-terminal sequence of a 75 kDa granule-bound protein (Rahman et al, 1995). This is illustrated in Figure 16. The location of the 75 kDa protein was  
5 determined for both the soluble fraction and starch granule-bound fraction by the method of Denyer et al (1995). Thus this cDNA clone encoded a polypeptide comprising a 41 amino acid transit peptide and a 606 amino acid mature peptide (SEQ ID NO:12). The cleavage site LRRL was located at amino  
10 acids 36 to 39 of the transit peptide of this deduced polypeptide.

Comparison of wheat SSS I with rice SSS and potato SSS showed that there is 87.4% or 75.9% homology at the amino acid level and 74.7% or 58.1% homology at the  
15 nucleotide level. Some amino acids in the at N-terminal sequences of the SSS I of wheat and rice were conserved. Major features of the SSS I gene are summarized in Table 3.

20 Example 19      Isolation of Genomic Clone of Wheat Soluble Starch Synthase

Seven genomic clones were obtained with a 300 bp cDNA probe by screening approximately  $5 \times 10^5$  plaques from a genomic DNA library of *Triticum tauschii*, as described above. DNA was purified from 5 of these clones and digested  
25 with *Bam*HI and *Sac*I. Southern hybridization analysis using the 300 bp cDNA as probe showed that these clones could be classified into two classes, as shown in Figure 17. One genomic clone, sg3, contained a long insert, and was digested with *Bam*HI or *Sac*I and subcloned into pBluescript  
30 KS+ vector.

- 41 -

**Table 3**  
**Comparison of exons and introns of soluble starch synthases**  
**I genes of wheat and rice**

(1) Identity of exons of soluble starch synthase I genes of wheat and rice

	Exons	wSSI-D1	rSSI	identity (%)	start site (wSSI-D1)	stop site (wSSI-D1)
	1a	255	113	57.52	-253	0
10	1b	316	298	58.92	1	316
	2	356	356	82.87	1473	1828
	3	78	78	92.31	2746	2823
	4	125	125	90.40	2906	3028
	5	82	82	89.02	4113	4194
15	6	174	174	93.10	4286	4459
	7	82	82	93.90	4562	4643
	8	92	92	92.39	4743	4835
	9	63	63	90.48	4959	5021
	10	90	90	82.22	5103	5192
20	11	125	125	88.80	8594	8718
	12	109	109	91.74	8807	8915
	13	53	53	81.13	8992	9044
	14	40	41	80.00	9160	9199
	15a	159	113	79.65	9499	9657
25	15b	392	539	46.46	9658	10098

(2) Identity of introns of soluble starch synthase I genes of wheat and rice

	Introns	wSSI-D1	rSSI	identity (%)	start site (wSSI-D1)	stop site (wSSI-D1)
	1	1156	907	41.05	317	1472
	2	917	851	41.65	1829	2745
	3	82	87	45.12	2824	2905
35	4	1084	835	48.50	3029	4112
	5	91	96	57.78	4195	4285
	6	102	189	52.48	4460	4561
	7	99	96	52.08	4644	4742
	8	123	110	45.46	4836	4958
40	9	81	78	58.97	5022	5102
	10	3401	663	37.56	5193	8593
	11	88	124	56.82	8719	8806
	12	76	81	48.68	8916	8991
	13	115	135	45.22	9045	9159
45	14	299	830	45.80	9200	9498

Note: Exon 1a: non-coding region of exon 1. Exon 1b: coding region of exon 1.

Exon 15a: coding region of exon 15. Exon 15b: non-coding region of exon 15.

50 wSSI-D1: wheat soluble starch synthase I gene.

rSSI: rice soluble starch synthase I gene.

- 42 -

These subclones were analysed by sequencing. The intron/exon structure of the sg3 rice gene is shown in Figure 18. The SSS I gene from *T. tauschii* is shown in SEQ ID No:13, while the deduced amino acid sequence is shown in  
5 SEQ ID NO:14.

Example 20                      Northern Hybridization Analysis of the  
Expression of Genes Encoding Soluble Starch  
Synthase

10                      Total RNAs were purified from leaves, pre-anthesis material, and various stages of developing endosperm at 5-8, 10-15 and 18-22 days post anthesis. Northern hybridization analysis showed that mRNAs encoding wheat SSS I were specifically expressed in developmental endosperm.  
15 Expression of this mRNAs in the leaves and pre-anthesis materials could not be detected by northern hybridization analysis under this experimental condition. Wheat SSS I mRNAs started to express at high levels at an early stage of endosperm, 5-8 days post anthesis, and the expression level  
20 in endosperm at 10-15 days post anthesis, was reduced. These results are summarized in Figure 9a and Figure 9d.

Example 21                      Genomic Localisation of Wheat Soluble Starch  
Synthase

25                      DNA from chromosome engineered lines was digested with the restriction enzyme BamHI and blotted onto supported nitrocellulose membranes. A probe prepared from the 3' end of the cDNA sequence, from positions 2345 to 2548, was used to hybridise to this DNA. The presence of a specific band  
30 was shown to be associated with the presence of chromosomes 7A (Figure 19). These data demonstrate location of the SSS I gene on chromosome 7.

Example 22                      Isolation of SSS I Promoter

35                      We have isolated the promoter that drives this pattern of expression for SSS I. The pattern of expression for SSS I is very similar to that for SBE II: the SSS I gene



transcript is detectable from an early stage of endosperm development until the endosperm matures. The sequence of this promoter is given in SEQ ID No:15.

5    Example 23                    Isolation of the Gene Encoding Debranching  
   Enzyme from Wheat

   The *sugary-1* mutation in maize results in mature dried kernels that have a glassy and translucent appearance; immature mature kernels accumulate sucrose and other simple  
10    sugars, as well as the water-soluble polysaccharide phytoglycogen (Black et al, 1966). Most data indicates that in *sugary-1* mutants the concentration of amylose is increased relative to that of amylopection. Analysis of a particular *sugary-1* mutation (*su-1Ref*) by James et al,  
15    (1995) led to the isolation of a cDNA that shared significant sequence identity with bacterial enzymes that hydrolyse the  $\alpha$  1,6-glucosyl linkages of starch, such as an isoamylase from *Pseudomonas* (Amemura et al, 1988), *ie.* bacterial debranching enzymes.

20                                    We have now isolated a sequence amplified from wheat endosperm cDNA using the polymerase chain reaction (PCR). This sequence is highly homologous to the sequence for the *sugary* gene isolated by James et al, (1995). This sequence has been used to isolate homologous cDNA sequences  
25    from a wheat endosperm library and genomic sequences from *Triticum tauschii*.

   Comparison of the deduced amino acid sequences of DBE from maize with spinach (Accession SOPULSPO, GenBank database), *Pseudomonas* (Amemura et al, 1988) and rice  
30    (Nakamura et al, 1997) enabled us to deduce sequences which could be useful in wheat. When these sequences were used as PCR amplification primers with wheat genomic DNA a product of 256 bp was produced. This was sequenced and was compared to the sequence of maize *sugary* isolated by James et al,  
35    (1995). The results are shown in Figure 20a and Figure 20b. This sequence has been termed wheat debranching enzyme sequence I (WDBE-I).

- 44 -

WDBE-1 was used to investigate a cDNA library constructed from wheat endosperm (Rahman et al, 1997) enables us to isolate two cDNA clones which hybridise strongly to the WDBE-I probe. The nucleotide sequence of the DNA insert in the longest of these clones is given in SEQ ID No:16.

Use of WDBE 1 to investigate a genomic library constructed from *T. tauschii*, as described above has led to the isolation of four genomic clones, designated I1, I2, I3 and I4, respectively, which hybridised strongly to the WDBE-I sequence. These clones were shown to contain copies of a single debranching enzyme gene. The sequence of one of these clones, I2, is given in SEQ ID No:17. The intron/exon structure of the gene is shown in Figure 20c. Exons 1 to 4 were identified by comparison with the maize *sugary-1* cDNA, while Exons 5 to 18 were identified by comparison with the cDNA sequence given in SEQ ID No:16. The major features of the DBE I gene are summarized in Table 4.

Hybridization of WDBE-I to DNA from *T. tauschii* indicates one hybridizing fragment (Figure 21a). The chromosomal location of the gene was shown to be on chromosome 7 through hybridisation to nullisomic/tetrasomic lines of the hexaploid wheat cultivar Chinese Spring (Figure 21b).

We have clearly isolated a sequence from the wheat genome that has high identity to the debranching enzyme cDNA of maize characterised by James et al (1997). The isolation of homologous cDNA sequences and genomic sequences enables further characterisation of the debranching enzyme cDNA and promoter sequences from wheat and *T. tauschii*. These sequences and the WDBE I sequences shown herein are useful in the manipulation of wheat starch structure through genetic manipulation and in the screening for mutants at the equivalent *sugary* locus in wheat.

Figure 9e shows that the DBE I gene is expressed during endosperm development in wheat and that the timing of expression is similar to the SBEII and SSSI genes. Figure 9h

- 45 -

shows that the full length mRNA for the gene (3.0 kb) is found only in the wheat endosperm.

Example 24      Transient assays of Promoter-GFP Fusions

5    **DNA constructs**

DNA constructs for transient expression assays were prepared by fusing sequences from the BEII and SSI promoters to the gene encoding the Green Fluorescent Protein. Green Fluorescent Protein (GFP) constructs contained the GFP gene described by Sheen et al. (1995). The nos 3' element (Bevan et al., 1983) was inserted 3' of the GFP gene. The plasmid vector (pWGEM\_NZfp) was constructed by inserting the NotI to HindIII fragment from the following sequence:

15    5' GCGGCCGCTC CCTGGCCGAC TTGGCCGAAG CTTGCATGCC TGCAGGTCGA  
CTCTAGAGGA TCCCCGGGTA CCGAGCTCGA ATTCATCGAT GATATCAGAT  
CCGGGCCCTC TAGATGCGGC CGCATGCATA AGCTT 3'

20    into the *NotI* and *HindIII* sites of pGem-13Zf(-) vector (Promega). The sequences at the junction of the wSSSIpro1 and wSSSIpro2 and GFP were identical, and included the junction sequence:

25    5'....CGCGCGCCCA CACCCTGCAG GTCGACTCTA GAGGATCCAT GGTGAGCAAG  
3'.

The sequence at the junction of wsbeIIpro1 and GFP was:

30    5' GCGACTGGCT GACTCAATCA CTACGCGGGG ATCCATGGTG AGCAAGGGCG  
3'.

The sequence at the junction of wsbeIIpro2 and GFP was:

35    5' GGACTCCTCT CGCGCCGTCC TGAGCCGCGG ATCCATGGTG AGCAAGGGCG  
3'.

The structures of the constructs are shown in Figures 22a to 22f.

- 46 -

**Table 4**  
**Structural features of wDBEI-D1**

**A.**  
**Position**  
**of exons**

Exon number	Start positi on	End posit ion	Comments
1	1890	2241	(deduced by comparison with maize)
2	2342	2524	(deduced by comparison with maize)
3	2615	2707	(deduced by comparison with maize)
4	3016	3168	(deduced by comparison with maize)
5	3360	3436	
6	4313	4454	
7	4526	4633	
8	4734	4819	
9	5058	5129	
10	5202	5328	
11	5558	5644	
12	6575	6671	
13	7507	7661	
14	8450	8527	
15	8739	8823	
16	8902	8981	
17	9114	9231	
18	Still being sequen ced		

- 5 Note that following nucleotides 3330, 6330 and 8419 there may be short regions of DNA not yet sequenced.

**B.**

	CAAAAT motif	1833
10	TCAAT motif	1838
	ATAAATAA motif	1804
	Endosperm box like motif TAAAACG	1463

- 47 -

**Preparation of target tissue**

All explants used for transient assay were from the hexaploid wheat cultivar, Milliwang. Endosperm (10 - 12 days after anthesis), embryos (12 - 14 days after anthesis) and leaves (the second leaf from the top of plants containing 5 leaves) were used. Developing seed or leaves were collected, surface sterilized with 1.25% w/v sodium hypochlorite for 20 minutes and rinsed with sterile distilled water 8 times. Endosperms or embryos were carefully excised from seed in order to avoid contamination with surrounding tissues. Leaves were cut into 0.5 cm x 1 cm pieces. All tissues were aseptically transferred onto SD1SM medium, which is an MS based medium containing 1 mg/L 2,4-D, 150 mg/L L-asparagine, 0.5 mg/L thiamine, 10 g/L sucrose, 36 g/L sorbitol and 36 g/L mannitol. Each agar plate contained either 12 endosperms, 12 embryos or 2 leaf segments.

**Preparation of gold particles and bombardment**

Five µg of each plasmid was used for the preparation of gold particles, as described by Witrzens et al. (1998). Gold particle-DNA suspension in ethanol (10 µl) was used for each bombardment using a Bio-Rad helium-driven particle delivery system, PDS-1000.

**GFP assay**

The expression of GFP was observed after 36 to 72 hours incubation using a fluorescence microscope. Two plates were bombarded for each construct. The numbers of expressing regions were recorded for each target tissue, and are summarized in Table 5. The intensity of the expression of GFP from each of the promoters was estimated by visual comparison of the light intensity emitted, and is summarized in Table 6.

The DNA construct containing GFP without a promoter region (pZLGFPNot) gave no evidence of transient expression in embryo (panel l) or leaf (panel r) and

- 48 -

extremely weak and sporadic expression in endosperm (panel f) , this construct gave only very weak expression in endosperm with respect to the number (Figure 5) and intensity (Figure 6) of transient expression regions. The constructs pwsssIpro1gfpNOT (panels b, h and n), psbeIIpro1gfpNOT(panels d, j and p), and psbeIIpro2gfpNOT (panels e, k and q) yielded low numbers (Table 5) of strongly (Table 6) expressing regions in leaves, and there was a very uneven distribution of expressing regions between target leaf pieces (Table 5). pwsssIpro2gfpNOT (panels c, i and o) gave no evidence of transient expression in leaves (Table 5). These results show that each of the promoter constructs is able to drive the transient expression of GFP in the grain tissues, endosperm and embryo. The ability of the short SSI promoter (pwsssIpro2gfpNOT containing 1042 bp 5' of the ATG translation start site) to drive expression in leaves (panel n) contrasts with the inability of the long SSI promoter (pwsssIpro2gfpNOT containing 3914 base pair region 5' of the ATG translation start site, panel o) ) suggesting that regions for controlling tissue specificity are located between -3914 and -1042 of the SSI promoter region (SEQ ID No:15).

#### Example 25      Stable transformation of rice

Stable transformation of rice using *Agrobacterium* was carried out essentially as described by Wang et al. 1997. The plasmids containing the target DNA constructs containing the promoter-reporter gene fusions are shown in Figure 23. These plasmids were transformed into *Agrobacterium tumefaciens* AGL1 by electroporation and cultured on selection plates of LB media containing rifampicillin (50 mg/L) and spectinomycin (50 mg/L) for 2 to 3 days, and then gently suspended in 10 ml NB liquid medium containing 100  $\mu$ M acetosyringone and mixed well. Embryogenic rice calli (2 to 3 months old) derived from mature seeds were immersed in the *A. tumefaciens* AGL1

**Table 5**  
**Transient Assay of GFP based constructs**

[illegible]

**Table 5 (Continued)**  
**Transient Assay of GFP based constructs**

[illegible]



**Table 5(Continued)**  
**Transient Assay of GFP based constructs**

[illegible]

- 52 -

**Table 6**  
**Comparison of the Intensities of Transient Expression**

Tissue	pact_j s- gfg_no s	pwsssI - prol <sub>gf</sub> pNOT	pwsssI - pro2 <sub>gf</sub> pNOT	psbeII - prol <sub>gf</sub> pNOT	psbeII - pro2 <sub>gf</sub> pNOT	pZLGFP Not
Endosperm	10	4	2.5	3.5	1.5	0.5
Embryo	10	5.5	5.5	1.5	1	0
Leaf	10	20	0	10	10	0

- 5 All intensities are relative to pact\_js-gfg\_nos transient expression in the target tissue  
 Relative intensities were independently scored by three researchers and averaged.

- 53 -

suspension. After 3 - 10 minutes the *A. tumefaciens* AGL1 suspension medium was removed, and the rice calli were transferred to NB medium containing 100  $\mu$ M acetosyringone for 48 h. The co-cultivated calli were washed with sterile Milli Q H<sub>2</sub>O containing 150 mg/L timentin 7 times to remove all *Agrobacterium*, plated on to NB medium containing 150 mg/L timentin and 30 mg/L hygromycin, and cultured for 3 to 4 weeks. Newly-formed buds on the surface of rice calli were excised and plated onto NB Second Selection medium containing 150 mg/L timentin and 50 mg/L hygromycin. After 4 weeks of proliferation calli were plated onto NB Pre-Regeneration medium containing 150 mg/L timentin and 50 mg/L hygromycin, and cultured for 2 weeks. The calli were then transferred on to NB-Regeneration medium containing 150 mg/L timentin and 50 mg/L hygromycin for 3 to 4 weeks. Once shooting occurs, shoots are transferred onto rooting medium (½ MS) containing 50 mg /L hygromycin. Once adequate root formation occurs, the seedlings are transferred to soil, grown in a misting chamber for 1-2 weeks, and grown to maturity in a containment glasshouse.

Example 26                      Use of probes from SSS I, SBE I, SBE II and DBE sequences to identify null or altered alleles for use in breeding programmes

DNA primer sets were designed to enable amplification of the first 9 introns of the SBE II gene using PCR. The design of the primer sets is illustrated in Figure 24. Primers were based on the wSBE II-D1 sequence (deduced from Figure 13b and Nair et al, 1997; Accession No. Y11282) and were designed such that intron sequences in the wSBE II sequence were amplified by PCR. These primer sets individually amplify the first 9 introns of SBE II. One primer (sr913F) contained a fluorescent label at the 5' end. Following amplification, the products were digested with the restriction enzyme DdeI and analysed using an ABI 377 DNA Sequencer with Genescan™ fragment analysis software. One primer set, for intron 5, was found to amplify products from

each of chromosomes 2A, 2B and 2D of wheat. This is shown in Figure 25, which illustrates results obtained with various wheat lines, and demonstrates that products from each of the wheat genomes from diverse wheats were amplified, and that therefore lines lacking the wSBEII gene on a specific chromosome could be readily identified. Lane (iii) illustrates the identification of the absence of the A genome wSBEII gene from the hexaploid wheat cultivar Chinese Spring ditelosomic line 2AS.

Figure 26 compares results of amplification with an Intron 10 primer set for various nullisomic/tetrasomic lines of the hexaploid wheat Chinese Spring. Fluorescent dUTP deoxynucleotides were included in the amplification reaction. Following amplification, the products were digested with the restriction enzyme *DdeI* and analysed using an ABI 377 DNA Sequencer with Genescan™ fragment analysis software. In lane (i) Chinese Spring ditelosomic line 2AS, a 300 base product is absent; in lane (ii) N2BT2A, a 204 base product is absent, and in lane (iii) N2DT2B a 191 base product is absent. These results demonstrate that the absence of specific wSBEII genes on each of the wheat chromosomes can be detected by this assay. Lines lacking wSBEII forms can be used as a parental line for breeding programmes for generation of new lines in which expression of SBE II is diminished or abolished, with consequent increase in amylose content of the wheat grain. Thus a high amylose wheat can be produced.

Table 7 shows examples primers pairs for SBE I, SSS I and DBE I which can identify genes from individual wheat genomes and could therefore be used to identify lines containing null or altered alleles. Such tests could be used to enable the development of wheat lines carrying null mutations in each of the genomes for a specific gene (for

Table 7  
PCR Primers for Starch Biosynthesis Genes

Gene	Forward Primer	Forward Primer sequence	Reverse Primer	Reverse Primer sequence	Temp (°C)	Product (bp)
SBE I	ZLE1 5d	GGC GGC GGC AAT GTG CGG CTG AG	ZLBE1 63	CCA GAT CGT ATA TCG GAA GGT CG	57.3	A=625, B = 600, D = 550
SSS I	ssSE01F	GAA CTC GCG CCC GAC CTC CT	ZLSg7	AGC CAC GAT TAT GCT GTC GAT GG	55.0	A, 450; B=450; D= 630
	ssSE14F	TTC TCA CCG CTA ACC GTG GAC	ZLSm19	GTC TAC ATG ACG TAG GGT TGG TC	55.8	B = 400, D = 500 no A product
DBE I	DBEE17F	TGG TCT GAG AAT AGC CGA TTC	sr1536F	AAGGCCACATAGATCTCG	56.8	B, 190; D, 190, A, 160. Non- specifi c product 220 bp

5 Temp: = annealing temperature, bp = length of the product in base pairs

- 56 -

example SBEI, SSI or DBE I) or combinations of null alleles for different genes.

It will be apparent to the person skilled in the art that while the invention has been described in some  
5 detail for the purposes of clarity and understanding, various modifications and alterations to the embodiments and methods described herein may be made without departing from the scope of the inventive concept disclosed in this specification.

10 Reference cited herein are listed on the following pages, and are incorporated herein by this reference.

REFERENCES

- Ainsworth, C., Clark, J. and Balsdon, J.  
Plant Molecular Biology, 1993 22 67-82
- 5 Amemura, A., Chakrabort, R., Fujita, M., Noumi, T. and  
Futai, M.  
Biol. Chem., 1988 263 9271-9275
- Baba, T., Kimura, K., Mizuno, K., Etoh, H., Ishida, Y.,  
Shida, O. and Arai, Y.
- 10 Biochem. Biophys. Res. Commun., 1991 181 87-94.  
Baba, T.; Nishihara, M.; Mizuno, K.; Kawasaki, T.; Shimada, H.;  
Kobayashi, E.; Ohnishi, S.; Tanaka, K.; Arai, Y.  
Plant Physiol, 1993, 103 565-573.
- Ball, S.; Guan, H.P.; James, M.; Myers, A.; Keeling, P.;
- 15 Mouille, G.; Buléon, A.; Colonna, P.; Preiss, J.  
Cell, 1996, 86 349-352
- Bevan, M., Barnes, W.M., and Chilton, M.  
Nucleic Acids Research, 1983, 11 369-385
- Black, R.C., Loerch, J.D., McARDle, F.J. and Creech, R.G.
- 20 Genetics, 1966 53 661-668
- Block, M., Loerz, H., Lutticke, S.  
Genbank database Accession number U48227
- Burton, R.A., Bewley, J.D., Smith, A.M., Bhattacharya, M.K.,  
Tatge, H., Ring, S., Bull, V., Hamilton, W.D.O. and Martin,  
25 C.  
The Plant Journal, 1995 7 3-15.
- Cangiano, G., La Volpe, A., Poulsen, P. and Kreiberg, J.D.  
Plant Physiology, 1993 102 1053-1054.
- Clarke, B.C., Mukai, Y. and Appels, R.
- 30 Chromosoma, 1996 105 269-275
- Deveraux, J., Haeberli, P. and Smithies, O.  
Nucleic Acids Res., 1984 12, 387-395.
- Denyer, K., Hylton, C.M., Jenner, C.F. and Smith, A.M.  
Planta, 1995 196 256-265
- 35 Doherty, J.P., Lindeman, R., Trent, R.J., Graham, M.W. and  
Woodcock, D.M.  
Gene, 1992 124 113-120

- Dry, I., Smith, A., Edwards, A., Bhattacharyya, M., Dunn, P., Martin, C.  
Plant J 1992, 2 193-202
- Edwards, A., Marshall, J., Sidebottom, C., Visser, R.G.F.,  
5 Smith, A.M., Martin, C.  
Plant J, 1995 8 283-294
- Fisher, D.K., Boyer, C.D. and Hannah, L.C.  
Plant Physiology, 1993 102 1045-1046
- Forde, B.G., Heyworth, A., Pywell, J. and Forde, M.  
10 Nucleic Acids Research, 1985 13 7327-7339
- Gill, B.S. and Appels, R.  
Plant Syst. Evol., 1988 160 77-90.
- Higgins, T.J.V., Zwar, J.A., Jacobsen, J.V. (1976)  
Nature, 1976, 260 166-168
- 15 Khandjian, E.W.  
Bio/Technology, 1987, 5 165-167
- Jahne, A., Lazzeri, P.A., Jager-Gussen, M. and Lorz, H.  
Theor. Appl. Genet., 1991 82 47-80
- James, M.G., Robertson, D.S. and Myers, A.M.  
20 Plant Cell, 1995 7 417-429
- Jolly, C.J., Glenn, G.M. and Rahman, S.  
Proc. Natl Acad. Sci., 1996 93 2408-2413.
- Kawasaki, T., Mizuno, K., Baba, T. and Shimada, H.  
Molec. Gen. Genet., 1993 237 10-16.
- 25 Lagudah, E.S., Appels, R. and McNeill, D.  
Genome, 1991 34 387-395
- Lazzeri, P.A., Brettschneider, R., Luhrs, R. and Lorz, H.  
Theor. Appl. Genet., 1991 81 437-444
- Maniatis, T., Fritsch, E.F. and Sambrook, J.  
30 Molecular cloning. A Laboratory Manual., New York. Cold  
Spring Harbor Laboratory, 1982
- Marshall, J.; Sidebottom, C.; Debet, M.; Martin, C.; Smith, A.M.;  
Edwards, A.  
The Plant Cell, 1996 8 1121-1135
- 35 Martin, C. and Smith, A.  
The Plant Cell, 1995 7 971-985.
- McElroy, D., Blowers, A.D., Jenes, B., Wu R.



- Mol. Gen. Genet., 1991 231 150-160.
- Mizuno, K., Kawasaki, T., Shimada, H., Satoh, H., Koyabashi, E., Okumura, S., Arai, Y. and Baba, T.
- J.Biol. Chem., 1993 268 19084-19091.
- 5 Muller, M.; Knudsen, S.
- Plant J, 1993, 4 343-355
- Morell, M.K., Blennow, A., Kosar-Hashemi, B. and Samuel, M.S.
- Plant Physiol., 1997 113 201-208.
- 10 Morell, M.K., Rahman, S., Abrahams, S.L. and Appels, R.
- Aust.J.of Plant Physiol., 1995 22 647-660.
- Nair, R., Baga, M., Scoles, G.J., Kartha, K. and Chibbar, R.
- Plant Science, 1997 122 153-163
- Nakamura, Y.; Kubo, A.; Shimamune, T.; Matsuda, T.; Harada, K.;
- 15 Satoh, H.
- Plant J, 1997, 12 143-153
- Nakamura, T., Yanamori, M., Hirano, H., Hidaka, S. and Nagamine, T.
- Molecular and General Genetics, 1995 248 253-259
- 20 Nakamura, Y., Takeichi, T., Kawaguchi, K. and Yamanouchi, H.
- Physiologia Plantarum, 1992 84 329-335.
- Nakamura, Y., Umemoto, T. and Sasaki, T.
- Planta, 1996 199 209-214
- Rahman, S., Kosar-Hashemi, B., Samuel, M., Hill, A., Abbott, D.C., Skerriett, J.H., Preiss, J., Appels, R. and Morell, M.
- 25 Aust. J. Plant Physiol., 1995 22 793-803.
- Rahman, S., Abrahams, S., Mukai, Y., Abbott, D., Samuel, M., Morell, M. and Appels, R.
- Genome, 1997 40 465-474
- 30 Repellin, A., Nair, R.B., Baga, M. and Chibbar, R.N.
- Plant Gene Register PGR97-094 (1997)
- Sambrook, J., Fritsch, E.F. and Maniatis, T.
- Molecular Cloning: A Laboratory Manual (Cold Spring Harbor Laboratory Press, 2<sup>nd</sup> ed 1989)
- 35 Sheen, J., Hwang, S., Niwa, Y., Kobayashi, H., and Galbraith, D.W.
- The Plant Journal, 1995 8 777-784

- Svensson, B.  
Plant Mol. Biol., 1994 25 141-157.  
Tanaka, K., Ohnishi, S., Kishimoto, N., Kawasaki, T., Baba, T.
- 5 Plant Physiol 1995, 108 677-683  
Tingay, S., McElroy, D., Kalla, R., Fieg, S., Wang, M.,  
Thornton, S. and Brettell, R.  
The Plant Journal, 1997 11 1369-1376  
Wan, Y. and Lemaux, P.G.
- 10 Plant Physiology, 1994 104 37-48  
Wang, M.B., Upadhyaya, N.M., Brettell, R.I.S., and  
Waterhouse, P.M.  
Journal of Genetics and Breeding, 1997 51 325-334.

- 61 -

## SEQUENCE LISTING

## (1) GENERAL INFORMATION:

## 5 (i) APPLICANT:

(A) NAME: COMMONWEALTH SCIENTIFIC AND INDUSTRIAL  
RESEARCH ORGANISATION

(B) STREET: Limestone Avenue

(C) CITY: Campbell

10 (D) STATE: ACT

(E) COUNTRY: AUSTRALIA

(F) POSTAL CODE (ZIP): 2612

(A) NAME: THE AUSTRALIAN NATIONAL UNIVERSITY

15 (B) STREET: BRIAN LEWIS CRESCENT

(C) CITY: ACTON

(D) STATE: ACT

(E) COUNTRY: AUSTRALIA

20 (F) POSTAL CODE (ZIP): 2601

(A) NAME: GOODMAN FIELDER LIMITED

(B) STREET: LEVEL 42, GROSVENOR PLACE

(C) CITY: SYDNEY

(D) STATE: NSW

25 (E) COUNTRY: AUSTRALIA

(F) POSTAL CODE (ZIP): 2000

(A) NAME: GROUPE LIMAGRAIN PACIFIC PTY LIMITED

30 (B) STREET: LEVEL 31, 1 O'CONNELL STREET

(C) CITY: SYDNEY

(D) STATE: NSW

(E) COUNTRY: AUSTRALIA

(F) POSTAL CODE (ZIP): 2000

35 (ii) TITLE OF INVENTION: REGULATION OF GENE EXPRESSION IN PLANTS

(iii) NUMBER OF SEQUENCES: 17

## (iv) COMPUTER READABLE FORM:

40 (A) MEDIUM TYPE: Floppy disk

(B) COMPUTER: IBM PC compatible

(C) OPERATING SYSTEM: PC-DOS/MS-DOS

(D) SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)

## 45 (2) INFORMATION FOR SEQ ID NO: 1:

## (i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 17 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

50 (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "pcr primer based on the N-terminal sequence of wSBE 15 ' end at  
position 168 of SEQ ID NO:5"

55

## (iii) HYPOTHETICAL: NO

- 62 -

- (iv) ANTI-SENSE:
- (v) FRAGMENT TYPE:
- 5 (vi) ORIGINAL SOURCE:  
(A) ORGANISM: triticum tauschii  
(F) TISSUE TYPE: Endosperm
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:
- 10 GGCACGCGAG AGACTGG 17
- (2) INFORMATION FOR SEQ ID NO: 2:
- (i) SEQUENCE CHARACTERISTICS:
- 15 (A) LENGTH: 19 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear
- 20 (ii) MOLECULE TYPE: other nucleic acid  
(A) DESCRIPTION: /desc = "pcr primer in which 5 ' end is at position 1590 of SEQ ID NO:5"
- (iii) HYPOTHETICAL: NO
- 25 (iv) ANTI-SENSE:
- (v) FRAGMENT TYPE:
- (vi) ORIGINAL SOURCE:
- 30 (A) ORGANISM: triticum tauschii  
(F) TISSUE TYPE: Endosperm
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2:
- 35 TACATTTTCCT TGTCCATCA 19
- (2) INFORMATION FOR SEQ ID NO: 3:
- (i) SEQUENCE CHARACTERISTICS:
- 40 (A) LENGTH: 18 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: other nucleic acid
- 45 (A) DESCRIPTION: /desc = "pcr primer 5 ' end is at position 1 of SEQ ID NO:5"
- (iii) HYPOTHETICAL: NO
- (iv) ANTI-SENSE:
- 50 (v) FRAGMENT TYPE:
- (vi) ORIGINAL SOURCE:
- 55 (A) ORGANISM: triticum tauschii  
(F) TISSUE TYPE: Endosperm

- 63 -

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 3:

ATCACGAGAG CTTGCTCA 18

5 (2) INFORMATION FOR SEQ ID NO: 4:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 23 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

10 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "per primer 5 ' end is at position 334 of SEQ ID NO:5"

15 (iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE:

(v) FRAGMENT TYPE:

20

(vi) ORIGINAL SOURCE:

(A) ORGANISM: triticum tauschii

(F) TISSUE TYPE: Endosperm

25 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 4:

CGGTACACAG TTGCGTCATT TTC 23

(2) INFORMATION FOR SEQ ID NO: 5:

30 (i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 2687 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

35 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

40 (iv) ANTI-SENSE:

(vi) ORIGINAL SOURCE:

(A) ORGANISM: triticum tauschii

(F) TISSUE TYPE: Endosperm

45

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 5:

ATCGACGAAG ATGCTCTGCC TCACCGCCCC CTCCTGCTCG CCATCTCTCC CGCCGCGCCC 60

50 CTCCCGTCCC GCTGCTGACC GGCCCGGACC GGGGATTTCG GCCAAGAGCA AGTTCTCTGT 120

TCCCGTGTCT GCGCCAAGAG ACTACACCAT GGCAACAGCT GAAGATGGTG TTGGCGACCT 180

55 TCCGATATAC GATCTGGATC CGAAGTTTGC CGGCTTCAAG GAACACTTCA GTTATAGGAT 240

GAAAAAGTAC CTTGACCAGA AACATTCGAT TGAGAAGCAC GAGGGAGGCC TTGAAGAGTT 300

CTCTAAAGGC TATTTGAAGT TTGGGATCAA CACAGAAAAT GACGCAACTG TGTACCGGGA 360

ATGGGCCCCCT GCAGCAATGG ATGCACAACCT TATTGGGTGAC TTCAACAACCT GGAATGGCTC 420  
 5 TGGGCACAGG ATGACAAAGG ATAATTATGG TGT TTGGTCA ATCAGGATTT CCCATGTCAA 480  
 TGGGAAACCT GCCATCCCCC ATAATTCCAA GGT TAAATTT CGATTT CACC GTGGAGATGG 540  
 ACTATGGGTC GATCGGGTTC CTGCATGGAT TCGTTATGCA ACTTTTGACG CCTCTAAATT 600  
 10 TGGAGCTCCA TATGACGGTG TTCACTGGGA TCCACCTTCT GGTGAAAGGT ATGTGTTTAA 660  
 GCATCCTCGG CCTCGAAAGC CTGACGCTCC ACGTATTTAC GAGGCTCATG TGGGGATGAG 720  
 15 TGGTGAGAGG CCTGAAGTAA GCACATACAG AGAATTTGCA GACAATGTGT TACCGCGCAT 780  
 AAAGGCAAAC AACTACAACA CAGTTCAGCT GATGGCAATC ATGGAACATT CCATATTATG 840  
 CTTCTTTTGG TACCATGTGA CGAATTTCTT CGCAGTTAGC AGCAGATCAG GAACACCAGA 900  
 20 GGACCTCAA TATCTTGTTG ACAAGGCACA TAGCTTAGGG TTGCGTGTTT TGATGGATGT 960  
 TGTCCATAGC CATGCGAGCA GTAATATGAC AGATGGTCTA AATGGCTATG ATGTTGGACA 1020  
 25 AAACACACAG GAGTCCTATT TCCATACAGG AGAAAGGGGT TATCATAAAC TGTGGGATAG 1080  
 TCGCCTGTTT AACTATGCCA ATTGGGAGGT CTTACGGTAT CTTCTTTCTA ATCTGAGATA 1140  
 TTGGATGGAC GAATTCATGT TTGACGGCTT CCGATTTGAT GGAGTAACAT CCATGCTATA 1200  
 30 TAATCACCAT GGTATCAATA TGTCATTTCG TGGAAATTAC AAGGAATATT TTGGTTTGGA 1260  
 TACCGATGTA GATGCAGTTG TTTACATGAT GCTTGCGAAC CATTTAATGC ACAAATCTT 1320  
 35 GCCAGAAGCA ACTGTTGTTG CAGAAGATGT TTCAGGCATG CCAGTGCTTT GTCGGTCAGT 1380  
 TGATGAAGGT GGAGTAGGGT TTGACTATCG CTTTGCTATG GCTATTCCTG ATAGATGGAT 1440  
 TGACTACTTG AAGAACAAAG ATGACCTTGA ATGGTCAATG AGTGCAATAG CACATACTCT 1500  
 40 GACCAACAGG AGATATACGG AAAAGTGCAT TGCATATGCT GAGAGCCACG ATCAGTCTAT 1560  
 TGT TGGCGAC AAGACTATGG CATTTCTCTT GATGGACAAG GAAATGTATA CTGGCATGTC 1620  
 45 AGACTTG CAG CCTGCTTCAC CTACAATTGA TCGTGAAT GCACCTCAA AGATGATTCA 1680  
 CTTCAATCACC ATGGCCCTTG GAGGTGATGG CTACTTGAAT TTTATGGGTA ATGAGTTTGG 1740  
 CCACCCAGAA TGGATTGACT TTCCAAGAGA AGGCAACAAC TGGAGTTATG ATAAATGCAG 1800  
 50 ACGCCAGTGG AGCCTCTCAG ACATTGATCA CCTACGATAC AAGTACATGA ACGCATTTGA 1860  
 TCAAGCAATG AATGCGCTCG ACGACAAGTT TTCCTTCCTA TCGTCATCAA AGCAGATTGT 1920  
 55 CAGCGACATG AATGAGGAAA AGAAGATTAT TGTATTTGAA CGTGGAGATC TGGTCTTCGT 1980  
 CTTCAATTTT CATCCCAGTA AAACCTATGA TGGTTACAAA GTCGGATGTG ATTTGCCTGG 2040  
 GAAGTACAAG GTAGCTCTGG ACTCCGATGC TCTGATGTTT GGTGGACATG GAAGAGTGGC 2100  
 60 CCAGTACAAC GATCACTTCA CGTCACCTGA AGGAGTACCA GGAGTACCTG AAACAACTT 2160  
 CAACAACCGC CCTAATTCAT TCAAAGTCCT GTCTCCACCC CGCACTTG TGCTTACTA 2220  
 65 TCGCGTCGAG GAAAAAGCGG AAAAGCCTAA GGATGAAGGA GCTGCTTCTT GGGGCAAAGC 2280  
 TGCTCCTGGG TACATCGATG TTGAAGCCAC TCGTGTCAA GACGCAGCAG ATGGTGAGGC 2340

- 65 -

GACTTCTGGT TCCAAAAAGG CGTCTACAGG AGGTGACTCC AGCAAGAAGG GAATTAACCTT 2400  
 TGTCTTCGGG TCACCTGACA AAGATAACAA ATAAGCACCA TATCAACGCT TGATCAGAAC 2460  
 5 CGTGTACCGA CGTCCTTGTA ATATTCCTGC TATTGCTAGT AGTAGCAATA CTGTCAAACCT 2520  
 GTGCAGACTT GAGATTCTGG CTTGGACTTT GCTGAGGTTA CCTACTATAT AGAAAGATAA 2580  
 10 ATAAGAGGTG ATGGTGCGGG TCGAGTCCGG CTATATGTGC CAAATATGCG CCATCCCGAG 2640  
 TCCTCTGTCA TAAAGGAAGT TTCGGGCTTT CAGCCCAGAA TAAAAAA 2687

## (2) INFORMATION FOR SEQ ID NO: 6:

## (i) SEQUENCE CHARACTERISTICS:

- 15 (A) LENGTH: 807 amino acids  
 (B) TYPE: amino acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

20 (ii) MOLECULE TYPE: protein

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE:

25

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: triticum tauschii  
 (F) TISSUE TYPE: Endosperm

30

(ix) FEATURE:

- (A) NAME/KEY: Protein  
 (B) LOCATION: 1..807  
 (D) OTHER INFORMATION: /label= sbel  
 /note= "deduced amino acid sequence from SEQ ID NO:5"

35

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 6:

40	Met	Leu	Cys	Leu	Thr	Ala	Pro	Ser	Cys	Ser	Pro	Ser	Leu	Pro	Pro	Arg
	1				5					10					15	
	Pro	Ser	Arg	Pro	Ala	Ala	Asp	Arg	Pro	Gly	Pro	Gly	Ile	Ser	Ala	Lys
				20					25					30		
45	Ser	Lys	Phe	Ser	Val	Pro	Val	Ser	Ala	Pro	Arg	Asp	Tyr	Thr	Met	Ala
			35					40					45			
	Thr	Ala	Glu	Asp	Gly	Val	Gly	Asp	Leu	Pro	Ile	Tyr	Asp	Leu	Asp	Pro
		50					55					60				
50	Lys	Phe	Ala	Gly	Phe	Lys	Glu	His	Phe	Ser	Tyr	Arg	Met	Lys	Lys	Tyr
	65					70					75				80	
	Leu	Asp	Gln	Lys	His	Ser	Ile	Glu	Lys	His	Glu	Gly	Gly	Leu	Glu	Glu
					85					90					95	
55	Phe	Ser	Lys	Gly	Tyr	Leu	Lys	Phe	Gly	Ile	Asn	Thr	Glu	Asn	Asp	Ala
				100					105					110		
60	Thr	Val	Tyr	Arg	Glu	Trp	Ala	Pro	Ala	Ala	Met	Asp	Ala	Gln	Leu	Ile
			115					120					125			

- 66 -

	Gly	Asp	Phe	Asn	Asn	Trp	Asn	Gly	Ser	Gly	His	Arg	Met	Thr	Lys	Asp
	130						135					140				
5	Asn	Tyr	Gly	Val	Trp	Ser	Ile	Arg	Ile	Ser	His	Val	Asn	Gly	Lys	Pro
	145					150					155					160
	Ala	Ile	Pro	His	Asn	Ser	Lys	Val	Lys	Phe	Arg	Phe	His	Arg	Gly	Asp
					165					170					175	
10	Gly	Leu	Trp	Val	Asp	Arg	Val	Pro	Ala	Trp	Ile	Arg	Tyr	Ala	Thr	Phe
				180					185					190		
	Asp	Ala	Ser	Lys	Phe	Gly	Ala	Pro	Tyr	Asp	Gly	Val	His	Trp	Asp	Pro
			195					200					205			
15	Pro	Ser	Gly	Glu	Arg	Tyr	Val	Phe	Lys	His	Pro	Arg	Pro	Arg	Lys	Pro
		210					215					220				
20	Asp	Ala	Pro	Arg	Ile	Tyr	Glu	Ala	His	Val	Gly	Met	Ser	Gly	Glu	Arg
	225					230					235					240
	Pro	Glu	Val	Ser	Thr	Tyr	Arg	Glu	Phe	Ala	Asp	Asn	Val	Leu	Pro	Arg
					245					250					255	
25	Ile	Lys	Ala	Asn	Asn	Tyr	Asn	Thr	Val	Gln	Leu	Met	Ala	Ile	Met	Glu
				260					265					270		
	His	Ser	Ile	Leu	Cys	Phe	Phe	Trp	Tyr	His	Val	Thr	Asn	Phe	Phe	Ala
			275					280					285			
30	Val	Ser	Ser	Arg	Ser	Gly	Thr	Pro	Glu	Asp	Leu	Lys	Tyr	Leu	Val	Asp
		290					295					300				
35	Lys	Ala	His	Ser	Leu	Gly	Leu	Arg	Val	Leu	Met	Asp	Val	Val	His	Ser
	305					310					315					320
	His	Ala	Ser	Ser	Asn	Met	Thr	Asp	Gly	Leu	Asn	Gly	Tyr	Asp	Val	Gly
					325					330					335	
40	Gln	Asn	Thr	Gln	Glu	Ser	Tyr	Phe	His	Thr	Gly	Glu	Arg	Gly	Tyr	His
				340					345					350		
	Lys	Leu	Trp	Asp	Ser	Arg	Leu	Phe	Asn	Tyr	Ala	Asn	Trp	Glu	Val	Leu
		355						360					365			
45	Arg	Tyr	Leu	Leu	Ser	Asn	Leu	Arg	Tyr	Trp	Met	Asp	Glu	Phe	Met	Phe
		370					375					380				
50	Asp	Gly	Phe	Arg	Phe	Asp	Gly	Val	Thr	Ser	Met	Leu	Tyr	Asn	His	His
	385					390					395					400
	Gly	Ile	Asn	Met	Ser	Phe	Ala	Gly	Asn	Tyr	Lys	Glu	Tyr	Phe	Gly	Leu
					405					410					415	
55	Asp	Thr	Asp	Val	Asp	Ala	Val	Val	Tyr	Met	Met	Leu	Ala	Asn	His	Leu
				420					425					430		
	Met	His	Lys	Ile	Leu	Pro	Glu	Ala	Thr	Val	Val	Ala	Glu	Asp	Val	Ser
			435					440					445			
60	Gly	Met	Pro	Val	Leu	Cys	Arg	Ser	Val	Asp	Glu	Gly	Gly	Val	Gly	Phe
		450					455					460				



- 67 -

	Asp	Tyr	Arg	Leu	Ala	Met	Ala	Ile	Pro	Asp	Arg	Trp	Ile	Asp	Tyr	Leu
	465					470					475					480
5	Lys	Asn	Lys	Asp	Asp	Leu	Glu	Trp	Ser	Met	Ser	Ala	Ile	Ala	His	Thr
				485						490					495	
	Leu	Thr	Asn	Arg	Arg	Tyr	Thr	Glu	Lys	Cys	Ile	Ala	Tyr	Ala	Glu	Ser
				500					505					510		
10	His	Asp	Gln	Ser	Ile	Val	Gly	Asp	Lys	Thr	Met	Ala	Phe	Leu	Leu	Met
			515					520					525			
	Asp	Lys	Glu	Met	Tyr	Thr	Gly	Met	Ser	Asp	Leu	Gln	Pro	Ala	Ser	Pro
15		530					535					540				
	Thr	Ile	Asp	Arg	Gly	Ile	Ala	Leu	Gln	Lys	Met	Ile	His	Phe	Ile	Thr
	545					550					555					560
20	Met	Ala	Leu	Gly	Gly	Asp	Gly	Tyr	Leu	Asn	Phe	Met	Gly	Asn	Glu	Phe
					565					570					575	
	Gly	His	Pro	Glu	Trp	Ile	Asp	Phe	Pro	Arg	Glu	Gly	Asn	Asn	Trp	Ser
				580					585					590		
25	Tyr	Asp	Lys	Cys	Arg	Arg	Gln	Trp	Ser	Leu	Ser	Asp	Ile	Asp	His	Leu
			595					600					605			
	Arg	Tyr	Lys	Tyr	Met	Asn	Ala	Phe	Asp	Gln	Ala	Met	Asn	Ala	Leu	Asp
30		610					615					620				
	Asp	Lys	Phe	Ser	Phe	Leu	Ser	Ser	Ser	Lys	Gln	Ile	Val	Ser	Asp	Met
	625					630					635					640
35	Asn	Glu	Glu	Lys	Lys	Ile	Ile	Val	Phe	Glu	Arg	Gly	Asp	Leu	Val	Phe
					645					650					655	
	Val	Phe	Asn	Phe	His	Pro	Ser	Lys	Thr	Tyr	Asp	Gly	Tyr	Lys	Val	Gly
				660					665					670		
40	Cys	Asp	Leu	Pro	Gly	Lys	Tyr	Lys	Val	Ala	Leu	Asp	Ser	Asp	Ala	Leu
			675					680					685			
	Met	Phe	Gly	Gly	His	Gly	Arg	Val	Ala	Gln	Tyr	Asn	Asp	His	Phe	Thr
45		690					695					700				
	Ser	Pro	Glu	Gly	Val	Pro	Gly	Val	Pro	Glu	Thr	Asn	Phe	Asn	Asn	Arg
	705					710					715					720
50	Pro	Asn	Ser	Phe	Lys	Val	Leu	Ser	Pro	Pro	Arg	Thr	Cys	Val	Ala	Tyr
					725					730					735	
	Tyr	Arg	Val	Glu	Glu	Lys	Ala	Glu	Lys	Pro	Lys	Asp	Glu	Gly	Ala	Ala
				740					745					750		
55	Ser	Trp	Gly	Lys	Ala	Ala	Pro	Gly	Tyr	Ile	Asp	Val	Glu	Ala	Thr	Arg
			755					760					765			
	Val	Lys	Asp	Ala	Ala	Asp	Gly	Glu	Ala	Thr	Ser	Gly	Ser	Lys	Lys	Ala
60		770					775					780				
	Ser	Thr	Gly	Gly	Asp	Ser	Ser	Lys	Lys	Gly	Ile	Asn	Phe	Val	Phe	Gly
	785					790					795					800

- 68 -

Ser Pro Asp Lys Asp Asn Lys  
805

- (2) INFORMATION FOR SEQ ID NO: 7:
- 5 (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 319 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear
- 10 (ii) MOLECULE TYPE: cDNA
- (iii) HYPOTHETICAL: NO
- 15 (iv) ANTI-SENSE:
- (vi) ORIGINAL SOURCE:  
(A) ORGANISM: triticum tauschii  
(F) TISSUE TYPE: Endosperm
- 20 (ix) FEATURE:  
(A) NAME/KEY: misc\_signal  
(B) LOCATION: 1..319  
(D) OTHER INFORMATION: /function= "3' untranslated region  
25 of wSBE I-D4 cDNA"
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 7:
- 30 GCGACTTCTG GTTCCAAAAA GGCGTCTACA GGAGGTGACT CCAGCAAGAA GGAATTAAC 60  
TTTGTCTTCG GGTCACCTGA CAAAGATAAC AAATAAGCAC CATATCAACG CTTGATCAGA 120  
ACCGTGTACC GACGTCCTTG TAATATTCCT GCTATTGCTA GTAGTAGCAA TACTGTCAAA 180  
35 CTGTGCAGAC TTGAGATTCT GGCTTGGACT TTGCTGAGGT TACCTACTAT ATAGAAAGAT 240  
AAATAAGAGG TGATGGTGCG GGTCGAGTCC GGCTATATGT GCCAAATATG CGCCATCCCCG 300  
40 AGTCCTCTGT CATAAAGGA 319
- (2) INFORMATION FOR SEQ ID NO: 8:
- (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 4890 base pairs  
(B) TYPE: nucleic acid  
45 (C) STRANDEDNESS: single  
(D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: DNA (genomic)
- 50 (iii) HYPOTHETICAL: NO
- (iv) ANTI-SENSE:
- (vi) ORIGINAL SOURCE:  
55 (A) ORGANISM: triticum tauschii  
(F) TISSUE TYPE: Endosperm
- (ix) FEATURE:

- 69 -

(A) NAME/KEY: promoter

(B) LOCATION: 1..4890

(D) OTHER INFORMATION: /function= "promoter containing  
sequence of SBE I"

5

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 8:

```

GGGTGGCGGG TCGGGCGGCA AGGCGCGGGG CGGCGGGGCG GCCGGGGCGG CGCGGCGGGC 60
10 CGGGCGGCAG CGGCGGCTAG GGTTCGCGG CGGCGGCGAC TTGGGCTGAG GCGGGGCACG 120
GGCTGCGGCT TTAAAGGCCG GCCAGGCTGA GGTGTCCGGG TCGGACACGG CCCGTAAGGC 180
15 GGTGACTTT AAAAATAAT AATTCGACA TGCAAAAAAG TAAGAAAAGA AATAATAAAC 240
GGACTCCAAA AATCCGAAG TAAATTTTTC CCCATTCTTA AAAATAAGCC GGACAAGATG 300
AACATTTATT TGGGCCTAAA ATGCAATTTT GAAAAATGCG TATTTTTCCT AATTCGGAAT 360
20 AAAATCAAAT AAAATCCAAA TAAAATCAA TATTTGTTTT TAATATTTTT CCTCCAATAT 420
TTCATTATTT GTGAAGAAGT CATTTTATCC CATCTCATAT ATTTTGATAT GAAATATTTT 480
25 CGGAGAGAAA AATAATTAAA ACAAATGATC CTATTTTCAA AATTGAGAA AACCCAAATA 540
TGAAAATAAC GAAATCCCA ACTCTCTCCG TGGGTCTTG AGTTGCGTGA AATTTCTAGG 600
ATCACAAATC AAAATGCAAT AAAATATGAT ATGCATGATG ATCTAATGTA TAACATTCCA 660
30 ATTGAAAATT TGGGATGTTA CATATAACTC AAATCTATA ATTATGAACA CAGAAATATT 720
AATGTAGAAC TCTATTTTGT TTTGAAATTG TATTATTTTT TAGAATTAGT CTAGAGCATT 780
TCGTGAACCT GAATCAAACC TTTAAATAAA ACAAAGCATA AAAATGACAA ATTCACATAT 840
35 GAAATAACTT GTGTTACATA GATTTATTAC AATAGCGTTG TATGTGTGTA TGTGTGCGTG 900
AGTGCCTATG GTAATATCAA TAAATATCTT GATAGATGTT TCTACAATTC ACGGGTCTAA 960
40 CTAGTAATGC AATGCAATGC ATGCTAAAAG AATAGAACCT TAGTTTCATT TAACTAACAA 1020
TTTTCAAATG TATGAGTTGC CAACAAGTGG CATACTTGGC ACTGTTTGTT TGTTCATTTT 1080
45 ATGGAAGTT CTTCTCTTTT TACATGGTTT AGATTCCAGC ATGTAGCCAC AAAATATGAT 1140
TGTCAAAAGA TAATACCTCA TAATACAATT CCACTAAAGT CACCTAGCCC AAGTGACCGA 1200
CCTGATCCTG AAATAAAATC AGAAGATTTG GTGTCATCAT CATGACAACA AATTATTAGG 1260
50 CGGTAGATCT TGTGGTAGTA CTCATGATGT AAAATTATCA AGAGGGAGAG AATGTATGGA 1320
GATTTATGTG AAGTACATCG TACACCAGAC ATAGTTGACA CATCGATTTT TTAAGATACA 1380
55 TTTGGACGCG CCTTGTGGGA GTGTAAAGTA CTACCATGTA TTAGAAGAGG TGAAATGAGA 1440
AATGCCATAG CTAGCAAGTA GGCCTAGTTA AGGAAATCTT TCCTTAGATC CCCTTCTCCC 1500
GAAGAGTGAA GTGCTTCAAC TAAAGGTTAG ACCCACTTAA AAAATGTCAC TTTGAATCTT 1560
60 TGCTTCCCTT GTCGTAATCC TGTGCATTTG TAGGTCCCTC GGATCTGAGC CCTTCTCCA 1620
AGCCCTTCAT TGGATTCCCC TGGATGTCTT TTTGTTACAT TTTATTGAAG TGAGAGTGAA 1680
TTATTATATG CCCATAGGAG GTGGGATATA AAGGCTGTTG GTATTCTGCA CCATACATGC 1740
65

```

- 70 -

TAGAGTAGGG AGGAGAGGCT GGTGCATGAT ACATGGTGGG CTAGCCCATA TATTTACCCC 1800  
TCCCCACCC ACTAACAAGT TTTTTTTATT AGGTCTTCAT CCTCTGATTT GTTTTTCTGT 1860  
5 TAGCCCATTC TTCATCATGG ACTTATTAAT CATGATTAGT TTCTTGGATT TTTGTTTACT 1920  
TGACTTGAAT TTGACAATGT GCCTCATATA TGGCATGTGG GACTGATAGG AAGATATATT 1980  
10 CTCACAACAT TAACTTAAAA AGGATTATTT TTTTGGTGCA GTCGTAAAGA AAACACTTTT 2040  
CTTTTATGCT AAAAGTTATT CAAACATAGA TTTATAAACA AAGGATATCA CCATGCATGA 2100  
CCATGCGCTC TCTCATGTTT ACTCTAGAAA CCATATATCT CTTTGTGCA AAATATTTAA 2160  
15 TCTATCCTCC TTGTTTCTGG GAATGAGTCG GGAAGGTAA TCTTAGGGAA GGTAAAGTG 2220  
AGGCAAGTAA GAGCAACTCT AGCAGAGTCG CGATATGCCC AATCGCCATA ATGCCAATAT 2280  
20 GGCATTTTGG GCCCAAATG GCACTTCAGA AGAGTCACCA TATCCCTTCG GATAGCCATA 2340  
ATTTAGGGAG CTCGCTCCAC AAACAAGCTT CGAGCCTCCA AATATGGAGG CCATGGATTC 2400  
GTTGTTTGGC ACTCACTCCA TATCCAACCG CAAGCGCATG CATGAGGGAA GTTTTAGCTT 2460  
25 CTTCTCCTT GCGCCAACGC CGGGATTTTA CACAGCGCAT TACAGGTACA TGAACCAGCA 2520  
TGCACAGATA ATCACCGACG AGTGGGGTGA CAAGAAGGAT AAGCACCTC CATTAGTGG 2580  
TGCGCCCACT CCCCTCAAAT TCATGAGGCA GCCATTTGGA TGGTCATCGC GTGGCATAAG 2640  
30 CTCCGACTAT AAAATCTCAA CGGCATCACC AAAACCATAG CTGCCGCTC CCCCTTCCTC 2700  
GGCATCACCT CCCCAAGACA TCTCCTCCCC TCTATGCCAC AATGTCATCA TTATGGAGAG 2760  
35 ACACAACTAC TGGTAAACCG CATACCCAAT CATGGTTTAC CGGCAGTGCG AACCCACCT 2820  
TCCTCCCACG ATGGTAGGAT ATTCTCCTCC TAGAATGGCG CGTGTGGCGC TTCCTCCTCC 2880  
CGAGGCTGAT ATGTCGGCTC CCATGATGGC GTGCATCATT GATTGTGGCG TCGGGTCCA 2940  
40 TCATACATGT TAACGAGGTC ATCCCATTG ATGTCGTTGG TCCCCTTGCC CCCAGTCGG 3000  
ATCCTGAGGA CCCGTTGAT GTCGCAATGC GACTCTCCAA ACTCAAAGCT CACAATGAGG 3060  
45 AGTACGTCCT CTAGGAGTTC CGCCCCGAA CCATCTATAA GGAGGAGCAA CGATAGCTCT 3120  
CCCCTACGCC TTCCTCGACG ATCTCTCTTA GGAGGACAAC GGCTAGACGA CGGCGGCGGC 3180  
GGCGAAGGTA CTGCAGGTAG TAGAACATAG CAATGTCGAA TGGCGACATT GCATATTTTG 3240  
50 AAAATGTCGC TCAACGACTT TTGAAGTCGC AAATAAAATG TAGTGTGACT ACTTTTGGCC 3300  
AGCAATATAA GTTTATCACA TTTGATAATG ATTTGAACCG GTGTGGTTCA ACTAAATGTA 3360  
55 CCATAAATTG AACATACAAA TTTTGTAGCA ATGAAAAAAG AAACAAGTAA GACCACAAAT 3420  
ATGAAAGCCG CATATCGCGA CTATGTGTTT GAGCCGCAGC TGCCAAGTAC ATATGAAGCG 3480  
TACTCCATAT GACATACGAC AACCATACAT ATGAAGACTC TACTAGAGTT CTCTAAGGCC 3540  
60 GCTTTTAGCG CCTTTCGTGC AGTGGTGCCC ATAGGGAGTG AGGGTAGTTG GACTGTTTCGT 3600  
TTCCCCTTTT TTCATTTCTT TGAAATCTAT TTTATTTTTT TTCTCTTTTG TAGGTTTCCC 3660  
65 AAATTTATAT ACCATTTTTC TGTTTCTCGC TATTTTTTGT TGTATATTC TAGTTTCATA 3720  
TTTTTCTATT ATTAATTTGT GTCTCTTATG AGAAGTCCAG ACTTGCATAT GGAGGTGCAC 3780

- 71 -

ACACAAACAT ATAAAGTATA AATACTAACT TGAGAAGTAT GTTTGCGTGG TCAAAAAAAC 3840  
 5 ATCATCAAAA CCTGCCAATA TGAGATATAG TTTTGAATAT ATCAATATGA GCAACGCAAC 3900  
 CATTTAAAAT GTGAACAATT GTTTTTTTTAG AAAAAATATA AGAAATAACT CCAACCCAGC 3960  
 CAAACCACAT GCTATACACT TGCTCCATAT GAAACCATGT TTGCTATTGG GCAGTTGCCT 4020  
 10 GAAACCGAAA GTAATGTTAG CCGTTTTTCT ATTCAAAGAA GAAGGAGAGT CGAGGTGACG 4080  
 CGATGCTTAG ACGTGAGATG GGGATGACCA CAACGTCCCT ACAGAGACCT CACCGGAGAT 4140  
 15 GGGGACATTG CAGTTGACAC GAGAGCGGTG AGGGGCTGCG ATGCGTGTGC GGCAACATGT 4200  
 GGCGAGGCGG ACGTCGGGCT GGCAGGTAGG GGGGAGGGGG AAGGACCGGG GGAGGAAGAA 4260  
 GAGGAGTAGC CTGCAAAACA TGGTACACCA GTTTTCTGCC CTACGAAAAC CTCATTTCAT 4320  
 20 TCCCCCACC TGACAAGCAA CAACCAACCA TCGCAGTCCC ACATGTCCCT CTGGTCTTTG 4380  
 CAAAAGTAA TTGTTCTTGC TGGACAGCGC AAAGAGTAAA CTTTGTGTTAG TTTTCATTTC 4440  
 25 TAGAAAAAGC AATCCTTTTA TAGTTCCTTTT GTGAAAGTAA TGCTTTTATA GTGATTGGGA 4500  
 TGTTCTTTTA GAGCAAATAT CTTCTTTTTT TTTTAGGGAA AAGAGCAAAT ATCTTCCACT 4560  
 TTTCACAAAA CTGACGAAGG CTGAAAGTGG CGAGACAGTG AGGGCCCCATA GCTTTCGTCC 4620  
 30 GGCCAGCGG CGCACGACCG TCCACGTGCA CCGCGGCCCT CCCGGGCCCCG CAGATCCGTT 4680  
 CTCCCTCGCC CCCGTTTCCC CCTCCCTCCC TCTCGTTGCT TCCACTCCAC TGTTCCTCCTC 4740  
 TTCCTGTCCA AAGCGGCCAC GGACCGGAAA AAAATCACGC CTTTCCGTTG GGTCTCCGGC 4800  
 35 GCCACACTCC TCCTCCGGCC GATATAAAGC GCGCGGGGCC ACGGGCCCCG CGCAAAATGG 4860  
 GATTCCCGTC CGCCGCCATG GAGGAAGATG 4890

40 (2) INFORMATION FOR SEQ ID NO: 9:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 6228 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 45 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

50

(iv) ANTI-SENSE:

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: triticum tauschii  
 55 (F) TISSUE TYPE: Endosperm

(ix) FEATURE:

- (A) NAME/KEY: misc\_feature  
 (B) LOCATION: 1  
 60 (D) OTHER INFORMATION: /product= "coding region of wSBE I-D4 gene"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 9:

- 72 -

	ACGGGCCCCG	CGCAAAATGG	GATTCCCGTC	CGCCGCCATC	GACGAAGATG	CTCTGCCTCA	60
	CCGCCCCCTC	CTGCTCGCCA	TCTCTCCCGC	CGCGCCCCCTC	CCGTCCCGCT	GCTGACCGGC	120
5	CCGGACCGGG	GATCTCGGTG	AGTCAGTCGG	GATCTTCATT	TCTTTTCTTT	TCTTTCGTTT	180
	CCGGCTCCGT	TCTGCCGGGG	TTTCCCTGAT	GCGATGCCGC	GCGCGCGCAG	GGCGGCGGCA	240
10	ATGTGCGGCT	GAGCGCGGTG	CCCGCGCCCT	CTTCGCTCCG	CTGGTCGTGG	CCGCGGAAGG	300
	TGAGCCCTCT	CCCCTGTCTA	CCCAGATTTG	CGACCGTGAT	CCCCTGTTGT	CGCCGGGCAA	360
	ACGGAATCTG	ATCCACGGTG	GTTATTGGAA	ATAGTATATA	CTACTAATAA	ACTTGAGGCT	420
15	GGGATTCGTC	CACTGAGGAA	CAAGTGGATG	CGATTTTCGAT	TGGATTTCTC	TGCTTTATGC	480
	GATCCGTACG	CAGAATATCC	CTCCTGCAGT	GTCTCAACCG	TATTACTGGA	TGTACAACCC	540
20	AAATGTGTAT	AATCTGTGCT	GAATGTATCA	ACCAATAATT	GCTGCATTGT	GAAAACATAA	600
	TCCTGTGTTG	TGTCTCTACT	ACTTGTTTCT	TCCTGATCTG	CCGCTTATCC	TAACCTTTGT	660
	TCATTTATGG	AAGGCAAGA	GCAAGTTCTC	TGTTCCCGTG	TCTGCGCCAA	GAGACTACAC	720
25	CATGGCAACA	GCTGAAGATG	GTGTTGGCGA	CCTTCCGATA	TACGATCTGG	ATCCGAAGTT	780
	TGCCGGGCTC	AAGGAACACT	TCAGTTATAG	GATGAAAAAG	TACCTTGACC	AGAAACATTC	840
30	GATTGAGAAG	CACGAGGGAG	GCCTTGAAGA	GTTCTCTAAA	GGTTAGCTTT	TGTTTCATGT	900
	GTTTGAAACA	ATAGTTACAT	CTTGTGGCGT	CCGCAGCACA	AAAGACATAA	TGCGACTCTG	960
	TTTTGTAGGC	TATTTGAAGT	TTGGGATCAA	CACAGAAAAT	GACGCAACTG	TGTACCGGGA	1020
35	ATGGGCCCCCT	GCAGCAATGT	AAGTTCTAGT	GTTGTCACGC	AACTAATTGC	AATGGTCGTT	1080
	GGTTAACTTA	TGAAGTGCTG	ATGAAACTGT	CTTAAGAGTT	TATGGCTTGT	CTTTTCTGAT	1140
40	TCTAGCTAGT	AAAGAGTAGA	TAAATATGAA	ATATGTTTTT	CCTTTTCTAG	TTATGGTCAT	1200
	GGTTGGCTGG	TATTCATTTT	TTTTATGGCA	ATACTTGCTT	CTAACTATCT	TTAGTAGATT	1260
	CATGTATTTA	CTTGTGAGTC	ATTACTTTAT	GGGTGTAGGG	ATGCACAAC	TATTGGTGAC	1320
45	TTCAACAAC	GGAATGGCTC	TGGGCACAGG	ATGACAAAGG	ATAATTATGG	TGTTTGGTCA	1380
	ATCAGGATTT	CCCATGTCAA	TGGGAAACCT	GCCATCCCCC	ATAATTCCAA	GGTTAAATTT	1440
50	CGATTCAC	GTGGAGATGG	ACTATGGGTC	GATCGGGTTC	CTGCATGGAT	TCGTTATGCA	1500
	ACTTTTGATG	CCTCTAAATT	TGGAGCTCCA	TATGACGGTG	TTCCTGCGGA	TCCACCTTCT	1560
	GGTGAAAGGT	CTACTTTTAG	TGGCTCGAGA	GCAAGAAATC	TAAGTAAAC	CCACACAATT	1620
55	AACTTACATT	AATGTGGAGA	CATGATACTT	TTATTGCTCG	TTTTGCAGGT	ATGTGTTTAA	1680
	GCATCCTCGG	CCTCGAAAGC	CTGACGCTCC	ACGTATTTAC	GAGGCTCATG	TGGGGATGAG	1740
60	TGGTGAAAAG	CCTGAAGTAA	GCACATACAG	AGAATTTGCA	GACAATGTGT	TACCGCGCAT	1800
	AAAGGCAAAC	AACTACAACA	CAGTTCAGCT	GATGGCAATC	ATGGAACATT	CATATTATGC	1860
	TTCTTTTGGG	TACCATGTGA	CGAATTTCTT	CGCAGTTAGC	AGCAGATCAG	AACGCCAGAG	1920
65	ACCTCAATAT	CTTGTTGACA	AGGCACATAG	TTTACGGTTG	CGTGTTCTGA	TGGATGTTGT	1980
	CCATAGCCAT	GCGAGCAGTA	ATAAGACAGA	TGGTCTTAAT	GGCTATGATG	TTGGGCAAAA	2040

	CACACAGGAG	TCCTATTTC	ACACAGGAGA	AAGGGGCTAT	CATAAACTGT	GGGATAGCCG	2100
5	CCTGTTCAAC	TATGCCAAT	GGGAGTCTTA	CGATTTCCTC	TTTCTAATCT	GAGATATTGG	2160
	ATGGACGAAT	TCATGTTTGA	TGGCTTCCGA	TTTGATGGGG	TAACATCCAT	GCTATATAAT	2220
	CACCATGGTA	TCAATATGTC	ATTGCTGGA	AGTTACAAGG	AATATTTTGG	TTTGGATACT	2280
10	GATGTAGATG	CAGTTGTTTA	CCTGATGCTT	GCGAACCATT	TAATGCACAA	ACTCTTGCCA	2340
	GAAGCAACTG	TTGTTGCAGA	AGATGTTTCA	GGCATGCCAG	TGCTTTGTCTG	GTCAGTTGAT	2400
15	GAAGGTGGAG	TAGGGTTTGA	CTATCGCCTG	GCTATGGCTA	TTCCTGATAG	ATGGATCGAC	2460
	TACTTGAAGA	ACAAAGATGA	CCTTGAATGG	TCAATGAGTG	GAATAGCACA	TACTCTGACC	2520
	AACAGGAGAT	ATACGGAAAA	GTGCATTGCA	TATGCTGAGA	GCCATGATCA	GGTATGTTTT	2580
20	CCCTCCTTTG	TCGCTGTGCG	TGAGTATGTG	TTCTTTTTTT	ATGGGGCACT	GGTCTAAGAA	2640
	CATACAGTTC	AAAGGTGAGA	CACTTTCTTT	GCCTGGTAGA	CAAATTTGAG	AAATAAACAT	2700
25	TTGCTTGAT	GACTTTTAGT	TGCTTCACAA	GTTTCAATTA	AGTTAGTTAT	ATTCTGATAA	2760
	CTAGTGATAG	TACCCACTAA	CCAGCTATTA	CGGACCATGT	AAGAATGTCC	GAAGACTGCA	2820
	GTTATATATC	GTTGACTTTG	TGTTCATCTA	TTGAAACAAC	TTAGTAGTTA	ACTTTCACGC	2880
30	AAATTTTCAG	TCTATTGTTG	GCGACAAGAC	TATGGCATT	CTCTTGATGG	ACAAGGAAAT	2940
	GTATACTGGC	ATGTCAGACT	TGCAGCCTGC	TTGCGCTACA	ATTGATCGTG	GAATTGCACT	3000
35	TCAAAAGGTT	CGATTCGTTT	TAAGTATTCC	TGAATTTGAT	GTTCTAGTTC	CAGACGAGTA	3060
	TTGTAATGTT	CGTTGTTACT	CAGAGTCTG	CTTAGTCCTT	GAAGATAATG	TATTCCAGTC	3120
	CCTTTTGGA	CATTTGGCTT	ATTTTGTTAC	AAATATTTCA	GATGATTCAC	TTCATCACCA	3180
40	TGGCCCTTGG	AGGTGATGGC	TACTTGAATT	TTATGGGTAA	TGAGGTAATA	TCTGGTTATC	3240
	TGTCAAAACT	TATTTCTGAT	CAATATGTTT	CGGGATTCCC	TCGAAAAAAA	TCCTTTGGGC	3300
45	AGGGCGAAAA	GTTTAAACAT	CTGTTTTCTA	TGATAGCCAA	GTACTCCCCA	GCTATTTCCA	3360
	TGTTATCACG	TATCATTTAG	CTGTGCCGGT	AGTTAATCTT	TATTCTAATT	CATTGTTGTT	3420
	TTTTAGCGTG	GCAGTCTATT	GTTGGATCCT	CTTATTCCAA	TTACATATAT	GCCGACATCA	3480
50	CACACTTATG	AATATTCCCT	GTTTAAAAGA	TTTTTATTTT	ATACCAATGT	TTCTCCGTAA	3540
	ATGATGCAAA	CATGATAGAG	ATGTTAGCAT	GTCTTTCTTA	ACCTACTCAT	GTTTTACATA	3600
55	TCACGACAAG	CTTCTTGCA	AAAATCAGCA	GTATATGGCA	AATTGCTGCA	ACCTGACAAC	3660
	GTTTATATCT	GTTTCTAAC	TCATACTGAC	GGTGCAATTT	CCTTTTAGTT	TGGCCACCCA	3720
	GAATGGATTG	ACTTTCAGA	AGAAGGCAAC	AAC TGGAGTT	ATGATAAATG	CAGACGCCAG	3780
60	TGGAGCCTCG	CAGACATTGA	TCACCTACGA	TACAAGGTTA	TGCCTATGTA	TATTTTTTACA	3840
	GTTTCTGGTC	TGGTAGCTCT	CTTGGGATCT	TGACCTCACT	TAGTTCCTTC	ATCTCTGACT	3900
65	GTAGCTTATT	TACACTGTGT	TCCAAC TTCT	GTCTTGTGGA	TAAATTCTCC	CTTCTAACGT	3960
	TTCATATTAA	GCCTTTTCAA	CTAAACTAAA	TTGCTGATCT	ACTACTAGTT	GCTCAGTACG	4020

- 74 -

ATGACCAAAT CTTGCCTGTG GTAACCTAGT AATTTTCTTG ATTCTTACAC ATTAGTGATA 4080  
TGCAGTGCAT ACATTATCCA TATAAATTGA CATTGCAATT TCCCAAATAT TATTTGAAGG 4140  
5 CTGTGTTCTT TTGTTAACAG GAAGTTATTT TCTCTGCATC TGATAAATAA TAATAGCCTT 4200  
TCACGATTTT TCTCATATTT TATCCAACCTT TTCTGCATTC AAGCATTTTTT TGTTCCTCGC 4260  
10 CTAACATATA TAATTTGAAC AGTACATGAA CGCATTTGAT CAAGCAATGA ATGCGCTCGA 4320  
CGACAAATTT TCCTTCCTAT CATCATCAAA GCAGATTGTC AGCGACATGA ATGAGGAAAA 4380  
GAAGTAGTTA ACTATACAAT GTTTAGTCAG GGCAGCTGTT GCATCATTTG ATTCACTCCT 4440  
15 ACTCTTAAGA ATAGCAACTC TGACTTGTGC GTTTTATGTT ACCAAATAAG TTGAAACCGT 4500  
ATCTGTTTGA TATGAACCAT TGTGTCTCA AAATGGGCTA TGGACTCAAT CCAACTTCCT 4560  
TTCCAGATTA TTGTATTTGA ACGTGGAATC TGGTCTTCGT CTTCAATTTT CATCCCAGTA 4620  
20 AAACCTTATGA TGGGTAACCTG ATCTCTTGCA AGCTTTGCCT TTCAATATTT CTTCTGCTTA 4680  
ATGACTAATG TGCTTAATCT CGTTTCCACT TTTAAAACAC GCAGTTACAA AGTCGGATGT 4740  
25 GACTTGCTG GGAAGTACAA GGTAAGCTCTG GACTCTGATG CTCTGATGTT TGGTGGACAT 4800  
GGAAGAGTAA GCAATGTTAA TGATGTTCAA GATCTGTTTT GCAACACTAT GTTCTTCTAT 4860  
AGAAGGGGCC ATCAAGGCTG CATCAGATAA TCTTATTTGC AGTGTTGATC TGTGCTGCAT 4920  
30 CGCAGGTGGC CCATGACAAC GATCACTTTA CGTCACCTGA AGGAGTACCA GGAGTACCTG 4980  
AAACAACTT CAACAACCGC CCTAACTCAT TCAAAATCCT GTCTCCATCC CGCACTTGTG 5040  
35 TGGTAATGCT AATTACTAGG AGGATTTAGT AACAATAAAT AAATAACAGC AAAAGATATC 5100  
TGCAGTACGA TCTCACAAA TGCTCTCTTG CCAGGCTTAC TATCGCGTCG AGGAGAAAGC 5160  
GGAAAAGCCC AAGGATGAAG GAGCTGCTTT CTTGGGGGAA ACTGCTCTCG GGTACATCGA 5220  
40 TGTGAAGCC ACTGGCGTCA AAGACGCAGC AGATGGTGAG GCGACTTCTG GTTCCGAAAA 5280  
GGCGTCTACA GGAGGTGACT CCAGCAAGAA GGAATTAAC TTTGTCTTTC TGTACCCGA 5340  
45 CAAAGACAAC AAATAAGCAC CATATCAACG CTTGATCAGG ACCGTGTGCC GACGTCCTTG 5400  
TAATACTCCT GCTATTGCTA GTAGTAGCAA TACTGTCAA CTGTGCAGAC TTGAAATTCT 5460  
GGCTTGACT TTGCTGAGGT TACCTACTAT ATAGAAAGAT AAATAAGCGG TGATGGTGCG 5520  
50 GGTGAGTCC AGCTATATGT GCCAAATATG CGCCATCCCG AGTCCTCTGT CATAAAGAAA 5580  
GTTTCGGGCT TCCATCCCAG AATAAAAACA GTTGTCTGTT TGCAATTTCT TTTTGTCTTG 5640  
55 CATAGTTACA TGATAATTGA TGCATATTGC TATAAGCCTG GATTGCATCT TCTTTTGCTA 5700  
ATAACTGCAG GGCCAAGAAA GCCTAGATTG TATCTTTTTT TGCTAATAAC TGCAGTGCTG 5760  
GGGAAGCTTC AGTCCTTGTT TCCGTTCTCG AGACAAGGCG TCATGTTTGG CGCACAAAGG 5820  
60 TAAGCCATCA TCTTATCAAG TCCCAAAAT CTCTGGTTGA AAGAAACCAT CACTAACTTG 5880  
TTCCAGGTGT TGGTTCCTCC ACAACCAAAA GCGGACCATC GTCGTCATCA TCGCTCACAG 5940  
65 CACTGACCAT CGAAGCCACG GTGGGCATGA AATGCGCATC GCCCAAGACT TGGGACCGTT 6000  
TCAAAATATC ACAAAGTACC ATGGCATCTT CTGCCAAAGG CTGCACTGCA CCTTTGGCAT 6060



5

10

(A) LENGTH: 11463 base pairs

(B) TYPE: nucleic acid

(D) TOPOLOGY: linear

15

(iii) **HYPOTHETICAL: NO**

## 20

(A) ORGANISM: *triticum tauschii*

25

(A) NAME/KEY: misc\_feature

(D) OTHER INFORMATION: /product= "complete sequence of the starch branching enzyme II gene"

30

## 35

CCAGTCCAAA ACGTCTGCTA GGATCACCAG CTGCAAAGTT AAGCGCGAGA CCACCAAAAC 180

40

ACAGCGGACG TGAGTGC GTG ACACATGGGG TCATCTATGG GCGTCGGAGC AAGGAAGAGA 300

45

ACATGCATTT TCAAACAAGG AAAATTAATT CTCAAACCAC CATGACATGC AATTCTCAA 480

50

GTCGAGTCGA GAAGAGGATG ACACTGAAAG TATGCGTATT ACGATTTCAT TTACATACAT 600

GTACAAATAC ATAATGTACC CTACAATTTG TTTTTTGGAG CAGAGTGGTG TGGTCTTTTT 660

55

TTTTTACACG AAAATGCCAT AGCTGGCCCG CATGCGTGCA GATCGGATGA TCGGTCGGAG 720

ACGACGGACA ATCAGACACT CACCAACTGC TTTTGTCTGG GACACAATAA ATGTTTTTTGT 780

60

AAACAAAATA AATACTTATA AACGAGGGTA CTAGAGGCCG CTAACGGCAT GGCCAGGTAA 840

ACGCGCTCCC AGCCGTTGGT TTGCGATCTC GTCCTCCCGC ACGCAGCGTC GCCTCCACCG 900

- 76 -

TCCGTCCGTC GCTGCCACCT CTGCTGTGCG CGCGCACGAA GGGAGGAAGA ACGAACGCCG 960  
 CACACACACT CACAACGGC ACACTCCCCG TGGGTCCCCT TTCCGGCTTG GCGTCTATCT 1020  
 5 CCTCTCCCCC GCCCATCCCC ATGCACTGCA CCGTACCCGC CAGCTTCCAC CCCCGCCGCA 1080  
 CACGTTGCTC CCCCTTCTCA TCGCTTCTCA ATTAATATCT CCATCACTCG GGTTCGCGC 1140  
 10 TGCATTTTCGG CCGGCGGGT GAGTGAGATC TGGGCGACTG GCTGACTCAA TCACTACGCG 1200  
 GGGATGGCGA CGTTCGCGGT GTCCGGCGCG ACTCTCGGTG TGGCGCGGGC CGGCGTCGGA 1260  
 GTGGCGCGGG CCGGCTCGGA GCGGAGGGGC GGGGCGGACT TGCCGTCGCT GCTCCTCAGG 1320  
 15 AAGAAGGACT CCTCTCGTAC GCCTCGCTCT CTCGAATCTC CCCCCTCTGG CTTTGGCTCC 1380  
 CCTTCTCTCT CCTCTGCGCG CGCATGGCCT GTTCGATGCT GTTCCCCAAT TGATCTCCAT 1440  
 GAGTGAGAGA GATAGCTGGA TTAGGCGATC GCGCTTCTTG AACCTGTATT TTTTCCCCCG 1500  
 20 CGGGGAAAATG CGTTAGTGTC ACCCAGGCCC TGGTGTACC ACGGCTTTGA TCATTCTCTCG 1560  
 TTTCAATTCTG ATATATATTT TCTCATTTCT TTTCTTCTTG TTCTTGCTGT AACTGCAAGT 1620  
 25 TGTGGCGTTT TTTCACATT GTAGTCATCC TTGCATTTTG CAGGCGCCGT CCTGAGCCGC 1680  
 GCGGCCCTCTC CAGGGAAGGT CCTGGTGCCT GACGGCGAGA GGACGACTTG GCAAGTCCGG 1740  
 CGCAACCTGA AGAATTACAG GTACACACAC TCGTGCCGGT AAATCTTCAT ACAATCGTTA 1800  
 30 TTCACTTACC AAATGCCGGA TGAAACCAAC CACGGATGCG TCAGGTTTCG AGCTTCTTCT 1860  
 ATCAGCATTG TGCAGTACTG CACTGCCTTG TTCATTTTGT TAGCCTTGGC CCCGTGCTGG 1920  
 35 CTCTTGGGCC ACTGAAAAA TCAGATGGAT GTGCATTCTA GCAAGAACTT CACAACATAA 1980  
 TGCACCGTTT GGGGTTTCGT CAGTCTGCTC TACAATTGCT ATTTTTCGTG CTGTAGATAC 2040  
 40 CTGAAGATAT CGAGGAGCAA ACGGCGGAAG TGAACATGAC AGGGGGGACT GCAGAGAAAC 2100  
 TTCAATCTTC AGAACCGACT CAGGGCATTG TGGAAACAAT CACTGATGGT GTAACCAAAG 2160  
 GAGTTAAGGA ACTAGTCGTG GGGGAGAAAC CGCGAGTTGT CCCAAAACCA GGAGATGGGC 2220  
 45 AGAAAAATATA CGAGATTGAC CCAACACTGA AAGATTTTCG GAGCCATCTT GACTACCGGT 2280  
 AATGCCTACC CGCTGCTTTC GCTCATTTTG AATTAAGGTC CTTTCATCAT GCAAATTTGG 2340  
 50 GGAACATCAA AGAGACAAAG ACTAGGGACC ACCATTTTCAT ACAGATCCCT TCGTGGTCTG 2400  
 AGAATATGCT GGGAAGTAAA TGTATAATTG ATGGCTACAA TTTGCTCAAA ATTGCAATAC 2460  
 GAATAACTGT CTCCGATCAT TACAATTAAA GAGTGGCAAA CTGATGAAAA TGTGGTGGAT 2520  
 55 GGGTTATAGA TTTTACTTTG CTAATTCCTC TACCAAATTC CTAGGGGGGA AATCTACCAG 2580  
 TTGGGAAACT TAGTTTCTTA TCTTTGTGGC CTTTTTGTTT TGGGGAAAC ACATTGCTAA 2640  
 60 ATTCTGAATGA TTTTGGGTAT ACCTCGGTGG ATTCAACAGA TACAGCGAAT ACAAGAGAAT 2700  
 TCGTGCTGCT ATTGACCAAC ATGAAGGTGG ATTGGAAGCA TTTTCTCGTG GTTATGAAAA 2760  
 GCTTGGATTT ACCCGCAGGT AAATTTAAAG CTTTATTATT ATGAAACGCC TCCACTAGTC 2820  
 65 TAATTGCATA TCTTATAAGA AAATTTATAA TTCCTGTTTT CCCCTCTCTT TTTTCCAGTG 2880  
 CTGAAGGTAT CGTCTAATTG CATATCTTAT AAGAAAATTT ATATTCCTGT TTTCCCCTAT 2940

TTTCCAGTGC TGAAGGTATC ACTTACCGAG AATGGGCTCC CTGGAGCGCA TGTTATGTTT 3000  
 5 TTTTAAGTTC CTTAACGAGA CACCTTCCAA TTTATTGTGA ATGGTCACTA TTCACCAACT 3060  
 AGCTTACTGG ACTTACAAAT TAGCTTACTG AATACTGACC AGTTACTATA AATTTATGAT 3120  
 CTGGCTTTTG CACCCGTGTA CAGTCTGCAG CATTAGTAGG TGACTTCAAC AATTGGAATC 3180  
 10 CAAATGCAGA TACTATGACC AGAGTATGTC TACAGCTTGG CAATTTTCCA CCTTTGCTTC 3240  
 ATAACTACTG ATACATCTAT TTGTATTTAT TTAGCTGTTT GCACATTCCT TAAAGTTGAG 3300  
 15 CCTCAACTAC ATCATATCAA AATGGTATAA TTTGTCAGTG TCTTAAGCTT CAGCCCAAAG 3360  
 ATTCTACTGA ATTTAGTCCA TCTTTTGTAG ATTGAAAATG AGTATATTAA GGATGAATGA 3420  
 ATACGTGCAA CACTCCCATC TGCATTATGT GTGCTTTTCC ATCTACAATG AGCATATTTT 3480  
 20 CATGCTATCA GTGAAGGTTT GCTCCTATTG ATGCAGATAT TTGATATGGT CTTTTCAGGA 3540  
 TGATTATGGT GTTTGGGAGA TTTTCCTCCC TAACAACGCT GATGGATCCT CAGCTATTCC 3600  
 TCATGGCTCA CGTGTAAGG TAAGCTGGCC AATTATTTAG TCGAGGATGT AGCATTTTCG 3660  
 25 AACTCTGCCT ACTAAGGGTC CTTTTTCCTC TCTGTTTTTT AGATACGGAT GGATACTCCA 3720  
 TCCGGTGTGA AGGATTCAAT TTCTGCTTGG ATCAAGTTCT CTGTGCAGGC TCCAGGTGAA 3780  
 30 ATACCTTTCA ATGGCATATA TTATGATCCA CCTGAAGAGG TAAGTATCGA TCTACATTAC 3840  
 ATTATTAAAT GAAATTTCCA GTGTACAGT TTTTAAATAC CCACTTCTTA CTGACATGTG 3900  
 35 AGTCAAGACA ATACTTTTGA ATTTGGAAGT GACATATGCA TTAATTCACC TTCTAAGGGC 3960  
 TAAGGGGCAA CCAACCTTGG TGATGTGTGT ATGCTTGTGT GTGACATAAG ATCTTATAGC 4020  
 TCTTTTATGT GTTCTCTGTT GGTTAGGATA TTCCATTTTG GCCTTTTGTG ACCATTTACT 4080  
 40 AAGGATATTT ACATGCAAAT GCAGGAGAAG TATGTCTTCC AACATCTCAA CTAAACGACC 4140  
 AGAGTCACTA AGGATTTATG AATCACACAT TGAATGAGC AGCCCGGTAT GTCAATAAGT 4200  
 TATTTACCT GTTCTGGTC TGATGGTTTA TTCTATGGAT TTTCTAGTTC TGTTATGTAC 4260  
 45 TGTAAACATA TTACATGGTG CATTCACTTG ACAACCTCGA TTTATTTTTC TAATGTCTTC 4320  
 ATATTGGCAA GTGCAAAACT TTGCTTCCTC TTTGTCTGCT TGTCTTTTGT TCTTCTGTAA 4380  
 50 GATTTCCATT GCATTTGGAG GCAGTGGGCA TGTGAAAGTC ATATCTATTT TTTTTTTGTC 4440  
 AGAGCATAGT TATATGAATT CCATTGTTGT TGCAATAGCT CGGTATAATG TAACCATGTT 4500  
 55 ACTAGCTTAA GATTTCCAC TTAGGATGTA AGAAATATTG CATTTGGAGCG TCTCCAGCAA 4560  
 GCCATTTCTT ACCTTATTAA TGAGAGAGAG ACAAGGGGGG GGGGGGGGGG GGGGTTCCCT 4620  
 TCATTATTCT GCGAGCGATT CAAAACTTC CATTGTTCTG AGGTGTACGT ACTGCAGGGA 4680  
 60 TCTCCCATTA TGAAGAGGAT ATAGTTAATT CTTTGTAACC TACTTGAAA CTTGAGTCTT 4740  
 GAGGCATCGC TAATATATAC TATCATCACA ATACTTAGAG GATGCATCTG AAATTTTAGT 4800  
 GTGATCTTGC ACAGGAACCG AAGATAAATT CATATGCTAA TTTTAGGGAT GAGGTGTTGC 4860  
 65 CAAGAATTAA AAGGCTTGGA TACAATGCAG TGCAGATAAT GGCAATCCAG GAGCATTCAT 4920

- 78 -

ACTATGCAAG CTTTGGGTAT TCACACAATC CATTTTTTTC TGTATACACT CTTCACCCAT 4980  
TTGGAGCTAT TACATCCTAA TGCTTCATGC ACATAAAATA TTTGGATATA ATCCTTTATT 5040  
5 AGATATATAG TACAACCTACA CTTAGTATTC TGAAAAAGAT CATTTTATTG TTGTTGGCTT 5100  
GTTCCAGGTA CCATGTTACT AATTTTTTTG CACCAAGTAG CCGTTTTTGA ACTCCAGAGG 5160  
10 ACTTAAAATC CTTGATCGAT AGAGCACATG AGCTTGGTTT GCTTGTTCTT ATGGATATTG 5220  
TTCATAGGTA ATTAGTCCAA TTTAATTTTA GCTGTTTTAC TGTTTATCTG GTATTCTAAA 5280  
GGGAAATTCA GGCAATTATG ATACATTGTC AAAAGCTAAG AGTGGCGAAA GTGAAATGTC 5340  
15 AAAATCTAGA GTGGCATAAG GAAAATTGGC AAAAAGTAGA GTGGCAAAAA TAAAATTTTC 5400  
CCATCCTAAA TGGCAGGGCC CTATCGCCGA ATATTTTTCC ATTCTATATA ATTGTGCTAC 5460  
GTGACTTCTT TTTTCTCAGA TGTATTAAAC CAGTTGGACA TGAAATGTAT TTGGTACATG 5520  
20 TAGTAAACTG ACAGTTCCAT AGAATATCGT TTTGTAATGG CAACACAATT TGATGCCATA 5580  
GATGTGGATT GAGAAGTTCA GATGCTATCA ATAGAATTAA TCAACTGGCC ATGTACTCGT 5640  
25 GGCACCTACAT ATAGTTTGCA AGTTGGAAAA CTGACAGCAA TACCTCACTG ATAAGTGGCC 5700  
AGGCCCCACT TGCCAGCTTC ATACTAGATG TTACTIONCCT GTTGAATTCA TTTGAACATA 5760  
TTACTTAAAG TTCTTCATTT GTCCTAAGTC AAACCTCTTT AAGTTTGACC AAGTCTATTG 5820  
30 GAAAATATAT CAACATCTAC AACACCAAAT TACTTTGATC AGATTAACAA TTTTTATTTT 5880  
ATTATATTAG CACATCTTTG ATGTTGTAGA TATCAGCACA TTTTCTATA GACTTGGTCA 5940  
35 AATATAGAGA AGTTTGACTT AGGACAAATC TAGAACTTCA ATCAATTTGG ATCAGAGGGA 6000  
ACATCAAATA ATATAGATAG ATGTCAACAC TTCAACAAAA AAATCAGACC TTGTCACCAT 6060  
ATATGCATCA GACCATCTGT TTGCTTTAGC CACTTGCTTT CATATTTATG TGTGTTGACC 6120  
40 TAATCTACTT TTCCTTCTAC TTGGTTTGGT TGATTCTATT TCAGTTGCAT TGCTTCATCA 6180  
ATGATTTTGT GTACCCTGCA GTCATTGCTC AAATAATACC CTTGACGGTT TGAATGGTTT 6240  
45 CGATGGCACT GATACACATT ACTTCCACGG TGGTCCACGC GGCCATCATT GGATGTGGGA 6300  
TTCTCGTCTA TTCAACTATG GGAGTTGGGA AGTATGTAGC TCTGACTTCT GTCACCATAT 6360  
TTGGCTAACT GTTCCTGTTA ATCTGTTCTT ACACATGTTG ATATTCCTATT CTTATGCAGG 6420  
50 TATTGAGATT CTTACTGTCA AACGCGAGAT GGTGGCTTGA AGAATATAAG TTTGATGGAT 6480  
TTCGATTTGA TGGGGTGACC TCCATGATGT ATACTCACCA TGGATTACAA GTAAGTCATC 6540  
55 AAGTGGTTTC AGTAACTTTT TTAGGGCACT GAAACAATTG CTATGCATCA TAACATGTAT 6600  
CATGATCAGG ACTTGTGCTA CGGAGTCTTA GATAGTTCCC TAGTATGCTT GTACAATTTT 6660  
60 ACCTGATGAG ATCATGGAAG ATTGGAAGTG ATTATTATTT ATTTTCTTTC TAAGTTTGTT 6720  
TCTTGTTCTA GATGACATTT ACTGGGAACT ATGGCGAATA TTTTGGATTT GCTACTGATG 6780  
TTGATGCGGT AGTTTACTTG ATGCTGGTCA ACGATCTAAT TCATGGACTT TATCCTGATG 6840  
65 CTGTATCCAT TGGTGAAGAT GTAAGTGCTT ACAGTATTTA TGATTTTTTAA CTAGTTAAGT 6900  
AGTTTTATTT TGGGGATCAG TCTGTTACAC TTTTGTGTTAG GGGTAAAATC TCTCTTTTCA 6960

TAACAATGCT AATTTATACC TTGTATGATA ATGCATCACT TAGTAATTTG AAAAGTGCAA 7020  
5 GGGCATTCAA GCTTACGAGC ATATTTTTTTG ATGGCTGTAA TTTATTTGAT AGTATGCTTG 7080  
TTTGGGTTTT TCAATAAGTG GGAGTGTGTG ACTAATGTTG TATTATTTAT TTAATTGCGG 7140  
AAGAAATGGG CAACCTTGTC AATTGCTTCA GAAGGCTAAC TTTGATTCCA TAAACGCTTT 7200  
10 GGAAATGAGA GGCTATTCCC AAGGACATGA ATTATACTTC AGTGTGTTCT GTACATGTAT 7260  
TTGTAATAGT GGTTTAACTT AAATTCCTGC ACTGCTATGG AATCTCACTG TATGTTGTAG 7320  
TGTACACATC CACAAACAAG TAATCCTGAG CTTTCAACTC ATGAGAAAAT AGAGTCCGCT 7380  
15 TCTGCCAGCA TTAAGTGTTC ACAGTTCTAA TTTGTGTAAC TGTGAAATTG TTCAGGTCAG 7440  
TGGAATGCCT ACATTTTGCA TCCCTGTTCC AGATGGTGGT GTTGGTTTTG ACTACCGCCT 7500  
20 GCATATGGCT GTAGCAGATA AATGGATTGA ACTCCTCAAG TAAGTGCAGG AATATTGGTG 7560  
ATTACATGCG CACAATGATC TAGATTACAT TTTCTAAATG GTAAAAAGGA AAATATGTAT 7620  
GTGAATATCT AGACATTTGC CTGTTATCAG CTTGAATACG AGAAGTCAAA TACATGATTT 7680  
25 AAATAGCAAA TCTCGGAAAT GTAATGGCTA GTGTCTTTAT GCTGGGCAGT GTACATTGCG 7740  
CTGTAGCAGG CCAGTCAACA CAGTTAGCAA TATTTTCAGA AACAATATTA TTTATATCCG 7800  
30 TATATGAGAA AGTTAGTATA TAAACTGTGG TCATTAATTG TGTTACCTT TTGTCCTGTT 7860  
TAAGGATGGG CAGTAGGTAA TAAATTTAGC CAGATAAAAT AAATCGTTAT TAGGTTTACA 7920  
AAAGGAATAT ACAGGGTCAT GTAGCATATC TAGTTGTAAT TAATGAAAAG GCTGACAAAA 7980  
35 GGCTCGGTAA AAAAACTTT ATGATGATCC AGATAGATAT GCAGGAACGC GACTAAAGCT 8040  
CAAATACTTA TTGCTACTAC ACAGCTGCCA ATCTGTCATG ATCTGTGTTT TGCTTTGTGC 8100  
40 TATTTAGATT TAAATACTAA CTCGATACAT TGGCAATAAT AAATTAATCT ATTCAACCAA 8160  
TTTGGTGGAT ACCAGAATTT CTGCCCTCTT GTTAGTAATG ATGTGCTCCC TGCTGCTGTT 8220  
CTCTGCCGTT ACAAAGCTG TTTTCAGTTT TTTGCATCAT TATTTTGTG TGTGAGTAGT 8280  
45 TTAAGCATGT TTTTGAAGC TGTGAGCTGT TGGTACTTAA TACATTCTTG GAAGTGTTCA 8340  
AATATGCTGC AGTGTAATTT AGCATTTCTT TAACACAGGC AAAGTGACGA ATCTTGAAA 8400  
50 ATGGGCGATA TTGTGCACAC CCTAACAAAT AGAAGGTGGC TTGAGAAGTG TGTAATTTAT 8460  
GCAGAAAGTC ATGATCAAGC ACTAGTTGGT GACAAGACTA TTGCATTCTG GTTGATGGAT 8520  
AAGGTACTAG CTGTTACTTT TGGACAAAAG AATTACTCCC TCCCGTTCCT AAATATAAGT 8580  
55 CTTTGTAGAG ATTCCACTAT GGACCACATA GTATATAGAT GCATTTTAGA GTGTAGATTC 8640  
ACTCATTTTG CTTCGTATGT AGTCCATAGT GAAATCTCTA CAGAGACTTA TATTTAGGAA 8700  
60 CGGAGGGAGT ACATAATTGA TTTGTCTCAT CAGATTGCTA GTGTTTCTT GTGATAAAGA 8760  
TTGGCTGCCT CACCCATCAC CAGCTATTTC CCAACTGTTA CTTGAGCAGA ATTTGCTGAA 8820  
AACGTACCAT GTGGTACTGT GGCGGCTTGT GAACTTTGAC AGTTATGTTG CAATTTTCTG 8880  
65 TTCTTATTTA TTTGATTGCT TATGTTACCG TTCATTTGCT CATTCCTTTC CGAGACCAGC 8940

	CAAAGTCACG	TGTTAGCTGT	GTGATCTGTT	ATCTGAATCT	TGAGCAAATT	TTATTAAATAG	9000
	GCTAAAATCC	AACGAATTAT	TTGCTTGAAT	TAAATATAC	AGACGTATAG	TCACCTGGCT	9060
5	CTTTCTTAGA	TGATTACCAT	AGTGCCTGAA	GGCTGAAATA	GTTTTGGTGT	TTCTTGGATG	9120
	CCGCCTAAAG	GAGTGATTTT	TATTGGATAG	ATTCCTGGCC	GAGTCTTCGT	TACAACATAA	9180
10	CATTTTGGAG	ATATGCTTAG	TAACAGCTCT	GGGAAGTTTG	GTCACAAGTC	TGCATCTACA	9240
	CGCTCCTTGA	GGTTTTATTA	TGGCGCCATC	TTTGTAACTA	GTGGCACCTG	TAAGGAAACA	9300
	CATTCAAAAG	GAAACGGTCA	CATCATTTCTA	ATCAGGACCA	CCATACTAAG	AGCAAGATTTC	9360
15	TGTTCCAATT	TTATGAGTTT	TTGGGACTCC	AAAGGGAACA	AAAGTGTCTC	ATATTGTGCT	9420
	TATAACTACA	GTTGTTTTTA	TACCAGTGTA	GTTTTATTCC	AGGACAGTTG	ATACTTGGTA	9480
20	CTGTGCTGTA	AATTATTTAT	CCGACATAGA	ACAGCATGAA	CATATCAAGC	TCTCTTTGTG	9540
	CAGGATATGT	ATGATTTTCAT	GGCTCTGGAT	AGGCTTCAAC	TCTTCGCATT	GATCGTGGCA	9600
	TAGCATTACA	TAAATGATC	AGGCTTGTC	CCATGGGTTT	AGGTGGTGAA	GGCTATCTTA	9660
25	ACTTCATGGG	AAATGAGTTT	GGGCATCCTG	GTCAGTCTTT	ACAACATTAT	TGCATTCTGC	9720
	ATGATTGTGA	TTTACTGTAA	TTTGAACCAT	GCTTTTCTTT	CACATTGTAT	GTATTATGTA	9780
30	ATCTGTTGCT	TCCAAGGAGG	AAGTTAACTT	CTATTTACTT	GGCAGAATGG	ATAGATTTTC	9840
	CAAGAGGCCC	ACAAACTCTT	CCAACCGGCA	AAGTTCTCCC	CTGGAAATAA	CAATAGTTAT	9900
	GATAAATGCC	GCCGTAGATT	TGATCTTGTA	AGTTTTAGCT	GTGCTATTAC	ATTCCCTCAC	9960
35	TAGATCTTTA	TTGGCCATTT	ATTTCTTGAT	GAAATCATAA	TGTTTGTTAG	GAAAGATCAA	10020
	CATTGCTTTT	GTAGTTTTGT	AGACGTTAAC	ATAAGTATGT	GTTGAGAGTT	GTTGATCATT	10080
40	AAAAATATCA	TGATTTTTTG	CAGGGAGATG	CAGATTTTCT	TAGATATCGT	GGTATGCAAG	10140
	AGTTCGATCA	GGCAATGCAG	CATCTTGAGG	AAAAATATGG	GGTATGTCAC	TGGTTTGTCT	10200
	TTGTTGCATA	ACAAGTCACA	GTTTAACGTC	AGTCTCTTCA	AGTGGTAAAA	AAAGTGTAGA	10260
45	ATTAATTCCT	GTAATGAGAT	GAAAACTGTG	CAAAGGCGGA	GCTGGAATTG	CTTTTCACCA	10320
	AAACTATTTT	CTTAAGTGCT	TGTGTATTGA	TACATATACC	AGCACTGACA	ATGTAAC'TGC	10380
50	AGTTTATGAC	ATCTGAGCAC	CAGTATGTTT	CACGGAAACA	TGAGGAAGAT	AAGGTGATCA	10440
	TCCTCAAAAG	AGGAGATTTG	GTATTTGTTT	TCAACTTCCA	CTGGAGCAAT	AGCTTTTTTTG	10500
	ACTACCGTGT	TGGGTGTTCC	AAGCCTGGGA	AGTACAAGGT	ATGCTTGCCCT	TTTCATTGTC	10560
55	CACCCTTCAC	CAGTAGGGTT	AGTGGGGGCT	TCTACAACTT	TTAATTCCAC	ATGGATAGAG	10620
	TTTGTTGGTC	GTGCAGCTAT	CAATATAAAG	AATAGGGTAA	TTTGTAAGA	AAAGAATTTG	10680
60	CTCGAGCTGT	TGTAGCCATA	GGAAGGTTGT	TCTTAACAGC	CCCGAAGCAC	ATACCATTC	10740
	TTCATATTAT	CTACTTAAGT	GTTTGTTTCA	ATCTTTATGC	TCAGTTGGAC	TCGGTCTAAT	10800
	ACTAGAACTA	TTTTCCGAAT	CTACCCTAAC	CATCCTAGCA	GTTTTAGAGC	AGCCCCATTT	10860
65	GGACAATTGG	CTGGGTTTTT	GTTAGTTGTG	ACAGTTTCTG	CTATTTCTTA	ATCAGGTGGC	10920
	CTTGGA'CTCT	GACGATGCAC	TCTTTGGTGG	ATTCAGCAGG	CTTGATCATG	ATGTCGACTA	10980

- 81 -

CTTCACAACC GTAAGTCTGG GCTCAAGCGT CACTTGACTC GTCTTGACTC AACTGCTTAC 11040  
 AAATCTGAAT CAACTTCCCA ATTGCTGATG CCCTTGACAGG AACATCCGCA TGACAACAGG 11100  
 5 CCGCGCTCTT TCTCGGTGTA CACTCCGAGC AGAACTGCGG TCGTGTATGC CCTTACAGAG 11160  
 TAAGAACCAG CAGCGGCTTG TTACAAGGCA AAGAGAGAAC TCCAGAGAGC TCGTGGATCG 11220  
 10 TGAGCGAAGC GACGGGCAAC GCGCGGAGGC TGCTCCAAGC GCCATGACTG GGAGGGGATC 11280  
 GTGCCTCTTC CCCAGATGCC AGGAGGAGCA GATGGATAGG TAGCTTGTTG GTGAGCGCTC 11340  
 GAAAGAAAAT GGACGGGCCT GGGTGTTTGT TGTGCTGCAC TGAACCCTCC TCCTATCTTG 11400  
 15 CACATTCCCG GTTGTTTTTG TACATATAAC TAATAATTGC CCGTGCGCTC AACGTGAAAA 11460  
 TCC 11463

## (2) INFORMATION FOR SEQ ID NO: 11:

## 20 (i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 2662 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

25

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

30 (iv) ANTI-SENSE:

(vi) ORIGINAL SOURCE:

(A) ORGANISM: triticum tauschii

(F) TISSUE TYPE: Endosperm

35

(ix) FEATURE:

(A) NAME/KEY: misc\_feature

(B) LOCATION: 1..2651

(D) OTHER INFORMATION: /product= "nucleotide sequence of  
 40 cDNA wheat SSS I"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 11:

45 TCTCCCACTC TTCTCTCCCC GCGCACACCG AGTCGGCACC GGCTCATCAC CCATCACCTC 60  
 GGCTTCGGCC ACCGGCAAAC CCCCCGATCC GCTTTTGCAG GCAGCGCACT AAAACCCCGG 120  
 GGAGCGCGCC CCGCGGCAGC AGCAGCACCG CAGTGGGAGA GAGAGGCTTC GCCCCGGCCC 180  
 50 GCACCGAGCG GGGCGATCCA CCGTCCGTGC GTCCGCACCT CCTCCGCCTC CTCCCCTGTC 240  
 CCGCGCGCCC ACACCCATGG CGGCGACGGG CGTCGGCGCC GGGTGCCTCG CCCCCAGCGT 300  
 CCGCCTGCGC GCCGATCCGG CGACGGCGGC CCGGGCGTCC GCCTGCGTCG TCCGCGCGCG 360  
 55 GCTCCGGCGC TTGGCGCGGG GCCGCTACGT TGCCGAGCTC AGCAGGGAGG GCCCGCGGGC 420  
 GCGCCCCGCG CAGCAGCAGC AACTGGCCCC GCCGCTCGTG CCAGGCTTCC TCGCGCCGCC 480  
 60 GCGCCCCGCG CCGGCCAGT CGCCGGCCCC GACGCAGCCG CCCCTGCCGG ACGCCGGCGT 540  
 GGGGGAATC GCGCCCGACC TCCTGCTCGA AGGGATTGCT GAGGATTCCA TCGACAGCAT 600

AATTGTGGCT GCAAGTGAGC AGGATTCTGA GATCATGGAT GCGAATGAGC AACCTCAAGC 660  
5 TAAAGTTACA CGTAGCATCG TGTTCGTGAC TGGTGAAGCT GCTCCTTATG CAAAGTCAGG 720  
GGGGCTGGGA GATGTTTGTG GTTCGTTACC AATTGCTCTT GCTGCTCGTG GTCACCGTGT 780  
GATGGTTGTA ATGCCAAGAT ACTTGAATGG GTCCTCTGAT AAAAATATG CAAAGGCATT 840  
10 ATACACTGGG AAGCACATTA AGATTCCATG CTTTGGGGGA TCACATGAAG TGACCTTTTTT 900  
TCATGAGTAT AGAGACAACG TCGATTGGGT GTTTGTGCGAT CATCCGTCAT ATCATAGACC 960  
15 AGGAAGTTTA TATGGAGATA ATTTTGGTGC TTTTGGTGAT AATCAGTTCA GATACACACT 1020  
CCTTTGCTAT GCTGCATGCG AGGCCCCACT AATCCTTGAA TTGGGAGGAT ATATTTATGG 1080  
ACAGAATTGC ATGTTTGTG TGAACGATTG GCATGCCAGC CTTGTGCCAG TCCTTCTTGC 1140  
20 TGCAAAATAT AGACCATACG GTGTTTACAG AGATTCCCGC AGCACCTTG TTATACATAA 1200  
TTTAGCACAT CAGGGTCTGG AGCCTGCAAG TACATATCCT GATCTGGGAT TGCCACCTGA 1260  
25 ATGGTATGGA GCTTTAGAAT GGGTATTTCC AGAATGGGCA AGGAGGCATG CCCTTGACAA 1320  
GGGTGAGGCA GTTAACCTTTT TGAAAGGAGC AGTCGTGACA GCAGATCGAA TTGTGACCGT 1380  
CAGTCAGGGT TATTCATGGG AGGTCACAAC TGCTGAAGGT GGACAGGGCC TCAATGAGCT 1440  
30 CTTAAGCTCC CGAAAAAGTG TATTGAATGG AATTGTAAAT GGAATTGACA TTAATGATTG 1500  
GAACCCACC ACAGACAAGT GTCTCCCTCA TCATTATTCT GTCGATGACC TCTCTGGAAG 1560  
GGCCAAATGT AAAGCTGAAT TGCAGAAGGA GCTGGGTTTA CCTGTAAGGG AGGATGTTCC 1620  
35 TCTGATTGGC TTTATTGGAA GACTGGATTA CCAGAAAGGC ATTGATCTCA TTAATGATGC 1680  
CATTCCAGAG CTCATGAGGG AGGACGTGCA GTTTGTCATG CTTGGATCTG GGGATCCAAT 1740  
40 TTTTGAAGGC TGGATGAGAT CTACCGAGTC GAGTTACAAG GATAAATTCC GTGGATGGGT 1800  
TGGATTTAGT GTTCCAGTTT CCCACAGAAT AACTGCAGGT TGCGATATAT TGTTAATGCC 1860  
ATCCAGGTTT GAACCTTGTG GTCTTAATCA GCTATATGCT ATGCAATATG GTACAGTTCC 1920  
45 TGTAAGTTCAT GGAAGTGGGG GCCTCCGAGA CACAGTCGAG ACCTTCAACC CTTTGGTGC 1980  
AAAAGGAGAG GAGGGTACAG GGTGGGCGTT CTCACCGCTA ACCGTGGACA AGATGTTGTG 2040  
50 GGCATTGCGA ACCGCGATGT CGACATTCAG GGAGCACAAG CCGTCCTGGG AGGGGCTCAT 2100  
GAAGCGAGGC ATGACGAAAG ACCATACGTG GGACCATGCC GCCGAGCAGT ACGAGCAGAT 2160  
CTTCGAATGG GCCTTCGTGG ACCAACCCTA CGTCATGTAG ACGGGGACTG GGGAGGTCGA 2220  
55 AGCGCGGGTC TCCTTGAGCT CTGAAGACAT GTTCCTCATC CTTCCGCGGC CCGGAAGGAT 2280  
ACCCCTGTAC ATTGCGTTGT CTTGCTACAG TAGAGTCGCA ATGCGCCTGC TTGCTTGGTC 2340  
60 CGCCGGTTTC AGAGTAGATG ACGGCTGTGC TGCTGCGGCG GTGACAGCTT CGGGTGGATG 2400  
ACAGTTACAG TTTTGGGGAA TAAGGAAGGG ATGTGCTGCA GGATGGTTAA CAGCAAAGCA 2460  
CCACTCAGAT GGCAGCCTCT CTGTCCGTGT TACAGCTGAA ATCAGAAACC AACTGGTGAC 2520  
65 TCTTTAGCCT TAGCGATTGT GAAGTTTGTT GCATTCTGTG TATGTTGTCT TGTCCTTAGC 2580



- 83 -

TGACAAATAT TAGACCTGTT GGAGAATTTT ATTTATCTTT GCTGCTGTTG TTTTGTTTT 2640  
 GTTAAAAAAA AAAAAA AAA 2662

## 5 (2) INFORMATION FOR SEQ ID NO: 12:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 768 amino acids  
 (B) TYPE: amino acid  
 (C) STRANDEDNESS: single  
 10 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(iii) HYPOTHETICAL: NO

15

## (vi) ORIGINAL SOURCE:

(A) ORGANISM: triticum tauschii

## (ix) FEATURE:

- 20 (A) NAME/KEY: Protein  
 (B) LOCATION: 1..768

## (ix) FEATURE:

- 25 (A) NAME/KEY: Protein  
 (B) LOCATION: 1..768  
 (D) OTHER INFORMATION: /product= "deduced amino acid  
 sequence SBE II"

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 12:

30

Met Ala Thr Phe Ala Val Ser Gly Ala Thr Leu Gly Val Ala Arg Pro  
 1 5 10 15

35

Pro Ala Ala Ala Gln Pro Glu Glu Leu Gln Ile Pro Glu Asp Ile Glu  
 20 25 30

Glu Gln Thr Ala Glu Val Asn Met Thr Gly Gly Thr Ala Glu Lys Leu  
 35 40 45

40

Glu Ser Ser Glu Pro Thr Gln Gly Ile Val Glu Thr Ile Thr Asp Gly  
 50 55 60

45

Val Thr Lys Gly Val Lys Glu Leu Val Val Gly Glu Lys Pro Arg Val  
 65 70 75 80

Val Pro Lys Pro Gly Asp Gly Gln Lys Ile Tyr Glu Ile Asp Pro Thr  
 85 90 95

50

Leu Lys Asp Phe Arg Ser His Leu Asp Tyr Arg Tyr Ser Glu Tyr Arg  
 100 105 110

Arg Ile Arg Ala Ala Ile Asp Gln His Glu Gly Gly Leu Glu Ala Phe  
 115 120 125

55

Ser Arg Gly Tyr Glu Lys Leu Gly Phe Thr Arg Ser Ala Glu Gly Ile  
 130 135 140

60

Thr Tyr Arg Glu Trp Ala Pro Gly Ala His Ser Ala Ala Leu Val Gly  
 145 150 155 160

- 84 -

	Asp	Phe	Asn	Asn	Trp	Asn	Pro	Asn	Ala	Asp	Thr	Met	Thr	Arg	Asp	Asp	
					165					170					175		
5	Tyr	Gly	Val	Trp	Glu	Ile	Phe	Leu	Pro	Asn	Asn	Ala	Asp	Gly	Ser	Pro	
				180					185					190			
	Ala	Ile	Pro	His	Gly	Ser	Arg	Val	Lys	Ile	Arg	Met	Asp	Thr	Pro	Ser	
			195					200					205				
10	Gly	Val	Lys	Asp	Ser	Ile	Ser	Ala	Trp	Ile	Lys	Phe	Ser	Val	Gln	Ala	
		210					215					220					
	Pro	Gly	Glu	Ile	Pro	Phe	Asn	Gly	Ile	Tyr	Tyr	Asp	Pro	Pro	Glu	Glu	
15		225				230					235					240	
	Glu	Lys	Tyr	Val	Phe	Gln	His	Pro	Gln	Pro	Lys	Arg	Pro	Glu	Ser	Leu	
				245					250						255		
20	Arg	Ile	Tyr	Glu	Ser	His	Ile	Gly	Met	Ser	Ser	Pro	Glu	Pro	Lys	Ile	
			260						265					270			
	Asn	Ser	Tyr	Ala	Asn	Phe	Arg	Asp	Glu	Val	Leu	Pro	Arg	Ile	Lys	Arg	
			275					280					285				
25	Leu	Gly	Tyr	Asn	Ala	Val	Gln	Ile	Met	Ala	Ile	Gln	Glu	His	Ser	Tyr	
		290				295						300					
	Tyr	Ala	Ser	Phe	Gly	Tyr	His	Val	Thr	Asn	Phe	Phe	Ala	Pro	Ser	Ser	
30		305				310					315					320	
	Arg	Phe	Gly	Thr	Pro	Glu	Asp	Leu	Lys	Ser	Leu	Ile	Asp	Arg	Ala	His	
				325						330					335		
35	Glu	Leu	Gly	Leu	Leu	Val	Leu	Met	Asp	Ile	Val	His	Ser	His	Ser	Ser	
			340						345					350			
	Asn	Asn	Thr	Leu	Asp	Gly	Leu	Asn	Gly	Phe	Asp	Gly	Thr	Asp	Thr	His	
			355					360					365				
40	Tyr	Phe	His	Gly	Gly	Pro	Arg	Gly	His	His	Trp	Met	Trp	Asp	Ser	Arg	
		370				375						380					
	Leu	Phe	Asn	Tyr	Gly	Ser	Trp	Glu	Val	Leu	Arg	Phe	Leu	Leu	Ser	Asn	
45		385				390					395					400	
	Ala	Arg	Trp	Trp	Leu	Glu	Glu	Tyr	Lys	Phe	Asp	Gly	Phe	Arg	Phe	Asp	
				405						410					415		
50	Gly	Val	Thr	Ser	Met	Met	Tyr	Thr	His	His	Gly	Leu	Gln	Met	Thr	Phe	
			420						425					430			
	Thr	Gly	Asn	Tyr	Gly	Glu	Tyr	Phe	Gly	Phe	Ala	Thr	Asp	Val	Asp	Ala	
			435					440					445				
55	Val	Val	Tyr	Leu	Met	Leu	Val	Asn	Asp	Leu	Ile	His	Gly	Leu	His	Pro	
		450				455						460					
	Asp	Ala	Val	Ser	Ile	Gly	Glu	Asp	Val	Ser	Gly	Met	Pro	Thr	Phe	Cys	
60		465				470					475					480	
	Ile	Pro	Val	Pro	Asp	Gly	Gly	Val	Gly	Phe	Asp	Tyr	Arg	Leu	His	Met	
				485						490						495	

- 85 -

	Ala	Val	Ala	Asp	Lys	Trp	Ile	Glu	Leu	Leu	Lys	Gln	Ser	Asp	Glu	Ser
				500					505					510		
5	Trp	Lys	Met	Gly	Asp	Ile	Val	His	Thr	Leu	Thr	Asn	Arg	Arg	Trp	Leu
			515					520					525			
	Glu	Lys	Cys	Val	Thr	Tyr	Ala	Glu	Ser	His	Asp	Gln	Ala	Leu	Val	Gly
			530				535					540				
10	Asp	Lys	Thr	Ile	Ala	Phe	Trp	Leu	Met	Asp	Lys	Asp	Met	Tyr	Asp	Phe
						550					555					560
	Met	Ala	Leu	Asp	Arg	Pro	Ser	Thr	Pro	Arg	Ile	Asp	Arg	Gly	Ile	Ala
15					565					570					575	
	Leu	His	Lys	Met	Ile	Arg	Leu	Val	Thr	Met	Gly	Leu	Gly	Gly	Glu	Gly
				580					585					590		
20	Tyr	Leu	Asn	Phe	Met	Gly	Asn	Glu	Phe	Gly	His	Pro	Glu	Trp	Ile	Asp
			595					600					605			
	Phe	Pro	Arg	Gly	Pro	Gln	Thr	Leu	Pro	Thr	Gly	Lys	Val	Leu	Pro	Gly
			610				615					620				
25	Asn	Asn	Asn	Ser	Tyr	Asp	Lys	Cys	Arg	Arg	Arg	Phe	Asp	Leu	Gly	Asp
						630					635					640
	Ala	Asp	Phe	Leu	Arg	Tyr	His	Gly	Met	Gln	Glu	Phe	Asp	Gln	Ala	Met
30					645					650					655	
	Gln	His	Leu	Glu	Glu	Lys	Tyr	Gly	Phe	Met	Thr	Ser	Glu	His	Gln	Tyr
				660					665					670		
35	Val	Ser	Arg	Lys	His	Glu	Glu	Asp	Lys	Val	Ile	Ile	Phe	Glu	Arg	Gly
			675					680					685			
	Asp	Leu	Val	Phe	Val	Phe	Asn	Phe	His	Trp	Ser	Asn	Ser	Phe	Phe	Asp
			690				695					700				
40	Tyr	Arg	Val	Gly	Cys	Ser	Arg	Pro	Gly	Lys	Tyr	Lys	Val	Ala	Leu	Asp
						710					715					720
	Ser	Asp	Asp	Ala	Leu	Phe	Gly	Gly	Phe	Ser	Arg	Leu	Asp	His	Asp	Val
45					725					730					735	
	Asp	Tyr	Phe	Thr	Thr	Glu	His	Pro	His	Asp	Asn	Arg	Pro	Arg	Ser	Phe
				740					745					750		
50	Ser	Val	Tyr	Thr	Pro	Ser	Arg	Thr	Ala	Val	Val	Tyr	Ala	Leu	Thr	Glu
			755					760					765			

## (2) INFORMATION FOR SEQ ID NO: 13:

## (i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 10550 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: both

(D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

- 86 -

- (vi) ORIGINAL SOURCE:  
(A) ORGANISM: triticum tauschii
- (ix) FEATURE:  
5 (A) NAME/KEY: exon  
(B) LOCATION:1..316  
(D) OTHER INFORMATION:/product= "exon 1"
- (ix) FEATURE:  
10 (A) NAME/KEY: exon  
(B) LOCATION:1472..1828  
(D) OTHER INFORMATION:/product= "exon 2"
- (ix) FEATURE:  
15 (A) NAME/KEY: exon  
(B) LOCATION:2766..2823  
(D) OTHER INFORMATION:/product= "exon 3"
- (ix) FEATURE:  
20 (A) NAME/KEY: exon  
(B) LOCATION:2906..3028  
(D) OTHER INFORMATION:/product= "exon 4"
- (ix) FEATURE:  
25 (A) NAME/KEY: exon  
(B) LOCATION:4113..4194  
(D) OTHER INFORMATION:/product= "exon 5"
- (ix) FEATURE:  
30 (A) NAME/KEY: exon  
(B) LOCATION:4286..4459  
(D) OTHER INFORMATION:/product= "exon 6"
- (ix) FEATURE:  
35 (A) NAME/KEY: exon  
(B) LOCATION:4562..4643  
(D) OTHER INFORMATION:/product= "exon 7"
- (ix) FEATURE:  
40 (A) NAME/KEY: exon  
(B) LOCATION:4744..4855  
(D) OTHER INFORMATION:/product= "exon 8"
- (ix) FEATURE:  
45 (A) NAME/KEY: exon  
(B) LOCATION:4999..5021  
(D) OTHER INFORMATION:/product= "exon 9"
- (ix) FEATURE:  
50 (A) NAME/KEY: exon  
(B) LOCATION:5102..5192  
(D) OTHER INFORMATION:/product= "exon 10"
- (ix) FEATURE:  
55 (A) NAME/KEY: exon  
(B) LOCATION:8593..8718

- 87 -

(D) OTHER INFORMATION:/product= "exon 11"

(ix) FEATURE:

(A) NAME/KEY: exon

5 (B) LOCATION:8807..8915

(D) OTHER INFORMATION:/product= "exon 12"

(ix) FEATURE:

(A) NAME/KEY: exon

10 (B) LOCATION:8992..9104

(D) OTHER INFORMATION:/product= "exon 13"

(ix) FEATURE:

(A) NAME/KEY: exon

15 (B) LOCATION:9161..9199

(D) OTHER INFORMATION:/product= "exon 14"

(ix) FEATURE:

(A) NAME/KEY: exon

20 (B) LOCATION:9498..9713

(D) OTHER INFORMATION:/product= "exon 15"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 13:

25	ATGGCGGCGA CGGGCGTCGG CGCCGGGTGC CTCGCCCCCA GCGTCCGCCT	50
	GCGCGCCGAT CCGGCGACGG CGGCCCGGGC GTCCGCTTGC GTCGTCCGCG	100
	CGCGGCTCCG GCGCTTGGCG CGGGGCCGCT ACGTCGCCGA GCTCAGCAGG	150
30	GAGGGCCCCG CGGCGCGCCC CGCGCAGCAG CAGCAACTGG CCCC GCCGCT	200
	CGTGCCAGGC TTCCTCGCGC CGCCGCCGCC CGCGCCCGCC CAGTCGCCGG	250
35	CCCCGACGCA GCCGCCCTG CCGGACGCCG GCGTGGGGGA ACTCGCGCCC	300
	GACCTCCTGC TCGAAGGTAA AAAACAAGGC TGAATCCTCA GATCACTCCG	350
	CGTCTTCGTT TTACCAAATA CGGTACTGCG AAGTGGTGCT GTATATGTGA	400
40	AGTTTCTGTC GATTTCTTCC TGACGGATGT TCAGTCGATT CAGTTGTATA	450
	TATGTGATAC GTTCGTTGTT CATCGATCGT ACAGATTTAC CAGCACACTA	500
45	GATAGAAATC GAGACCGACG CGGGCAGATC AATAGATTTT TCTAGACGTT	550
	TTATTGGATC GTGAGATGAT TGATTGGGGT GGCCTGTCTGA TACGATAGCG	600
	GTGCACCGCC GATGTATCGG GGCATGTGCA CGTGGTTGGG TCTCAGCAGA	650
50	CATATCACTA GACTGGTATC GTAATTTACT AGTACTACTG GAAAGAGGAC	700
	TAAAAAGGCT AGGCCAAGTG CACGCATGTT GGGAACGTTG TTAAATTGAT	750
55	GAGTTTGTCC TTTGCTTGGG CTGGTATTAT TACCAAAAAA TGGTGTTAGT	800

	CCCTGTACTT ATTAATGGGA AAATCTTAAC ATGACACTGG GGTTTATGAG	850
	TCTCCAATTG TATATTCTCA GCACTCAACT GATTTTACTG ATACTGTAGT	900
5	GGAAATGACA CGTGAGCACC CCCCTTCAAG GAATGCAATG CTTCCTTCTG	950
	TTTTATATTA CAGGAAGTAG AAGGAGCTTC CACCTTTGAG TACAGAAGTA	1000
10	CTCCCTCCGT TCCAAAATAG ATGACTCAAC TTTGTACTAA TTTGTACTA	1050
	TAGTTAGTAC AAAGTTGAGT CATCTATTTT AGAACGGAGG GAGTAGTATC	1100
	GAAATTGAAG ACCCTTGTAT TACTGTCTTG TTTTCAATG AAAATGGGAG	1150
15	GCCCATGCAG TAAGTCACAT GGGCACCTGG GAGGCTGGGA TCATGTGTGC	1200
	TTTGCAGAGT ACTAGACCCA GCTCACCTC TGTTAGATTA CTTGTTGGGC	1250
	TGCTACTTTG TGTTTGCTGT GCAGTATATC AGACATCCTG AATTTGGCAT	1300
20	CTAGCTGAGA ACAGAATGCA GGTTGCACCA TTCTTATTAT TGCTAAACTG	1350
	TTGTCACGCA ATTTATAAAG AATGTGATCT TCTGAGTATT AATTAATCAT	1400
25	GTTCTGCTAA TATCTGTCCT CGCTCTGGTG TTGACAAATA TACCATATGA	1450
	ATATTTTCCA TTTTGCAACC AGGGATTGCT GAGGATTCCA TCGACAGCAT	1500
	AATCGTGGCT GCAAGTGAGC AGGATTCTGA GATCATGGAT GCGAATGAGC	1550
30	AACCTCAAGC TAAAGTTACA CGTAGCATCG TGTTTGTGAC TGGTGAAGCT	1600
	GCTCCTTATG CAAAGTCAGG GGGGCTGGGA GATGTTTGTG GTTCGTTACC	1650
35	AATTGCTCTT GCTGCTCGTG GTCACCGTGT GATGGTTGTA ATGCCAAGAT	1700
	ACTTGAATGG GTCCTCTGAT AAAAATATG CAAAGGCATT ATACACTGCG	1750
	AAGCACATTA AGATTCCATG CTTTGGGGGA TCACATGAAG TGACCTTTTT	1800
40	TCATGAGTAT AGAGACAACG TCGATTGGGT GGGTACACAA TCACCTTCTT	1850
	ATTCTCTGTT GAATTGTAGC AACTGTTTAT CCTTGTTTAC ACTTCTTTTA	1900
45	GCCCTGCAAA GACATATGTG ATTTCCATAC TTTTTTGTTA TTTCCCTGT	1950
	ACTCTTGCTC ATGAAGGTCA AAATATCATA TATCCATGGA AGTCATGCAT	2000
	GTGCCTAGTA TTTTGGTGT CGGTGCCTTT AACTTTCAGG GATTAATACG	2050
50	TGGAATTTGA TAACTAAAGT TTATTTTATT GAAAAAATT GTAGGTGG	2100
	TGAGCCCACA GCCACGCAGT GGCACCACTG CTTGCACATG ATTTTGCATT	2150
55	TCTGTTTGCA CCGAGCACTT CATGTGAATA AGGTGTAAAA TCATAAAGTA	2200

- 89 -

	CCAATTTTAT TCTGCCAATT GCACTTAAGA GTATATACAT TTATCTTGGC	2250
	CTCAATCATG GGAGTACTGT GCATTCAGTG CACCATCATT GTTCTAAGGA	2300
5	GAAAATGTGG GTGCAAGGAA GACACTTTTG TCCCTTAATA AAAGGCAGGC	2350
	ACTCTGTTGT CATATAGATA GAAAGCAACA AACTTATTTC AAAGAGCTAA	2400
10	CAATGGCAAA AGAACCACAAA AAAGCATGCT AAGGCGGTGA CACCAAAAGG	2450
	TGAGGGGGGC CTTGTGACTG ACAGCACCCC AAACCTATTGC CATTGTTTTA	2500
	CTAAATGAAG ATCATTTTAG AAGCTCTCAG GAACTTCGAA AACAGTGGCT	2550
15	TTCCGTCCAC AGATCGTCTG TTAATATTTT TGTCCAGTGA TACTTTTTTT	2600
	GCTCCTTACA AGAGTGCCTA TGTTGACATA TACATTGTTA AGTTGTTCAT	2650
20	AAGTTTACTT CTTATTCTAA ACAGCAAGTG CCTAATGCTT GCATTTATTT	2700
	TGGCTATTTA TTTTATTCT CATTCAATC AACACTTTTG TTCAGGTGTT	2750
	TGTCGATCAT CCGTCATATC ATAGACCAGG AAGTTTATAT GGAGATAATT	2800
25	TTGGTGCTTT TGGTGATAAT CAGGTACACT AACTATACT AAGCTCCTAG	2850
	TTGACTAAGT CGTAAGTTGT ACCTCCTCGC TGACCGGCTG CTCTATGTCG	2900
30	TGCAGTTCAG ATACACACTC CTTTGCTATG CTGCATGCGA GGCCCCACTA	2950
	ATCCTTGAAT TGGGAGGATA TATTTATGGA CAGAATTGCA TGTTTGTTGT	3000
	GAACGATTGG CATGCCAGCC TTGTGCCAGT GTACGTTGTT TGTGGATCTG	3050
35	AAAGTCCAAT CCTTTATTCA TTCTCTGCTT TGCAGTGTGC CCATGTCTAC	3100
	ATTTCTTTTA TGCTTTTTTC ATGTCTGTTT TTATATTGCA TATATGCTTA	3150
40	TGGAGTCTAA AAGTTACCGG AGGGAATAAC TCTTAAGGAT TTCCTCAATC	3200
	AATTATCTTT AGCTTTAGTT AACATTTACT GTGGCAAACA TAATGTGTTT	3250
	TGAGATTTAC AAGTTCAGAG ATTGCACTTC ACTAGTTCGT AGCTAATCTG	3300
45	ATGTTTTCCC CGAGAAAATG CCTAAAGCTT TGTGTCTTGA TGCATTGATA	3350
	GAAAAAGAGT TTATGTACAC TCCCAAAGAG GGGACCCAAA ATTACAACAC	3400
50	CACACCCCTG AGAACTAGGC GCTGCCGGAA GAAGCGATGC AAGCCCCACT	3450
	GCCCCTGCCT TAGCTCAAAG CCGGGCGTCA GCTTGATTGT GTCAAGTAAG	3500
	CTAGCAGTGC TAGATTGCGC AAGGTGCGATT CGTCGAAGAT GACAGTGTTG	3550
55	CGCTGCTTCC AAATCCACCA AACTATGAGC ATGATCACTG GAGAAGTACC	3600

- 90 -

	TTTTCTCGCG GCTGAGGGGG TGGACTGGTG GTCTGCTGCT GCCAGTTTTC	3650
	AGATAATCTG AAAAATGCAT GTTTTGATGA TTTTAGTATC TTGCGGACCC	3700
5	TGGGTACCAC CTAAGCTTTC ACACAGTAAT TTGCAGTTAC ACCTATAAAA	3750
	GTAACGGTCA TGATATGCAT GTGTTTTGGG TAGATCATGG TGCATGCATT	3800
	TTAGGAATTA GGACATGCCA GAACCACGTG AGGCTTATGG GGCAATTTCAT	3850
10	TTGTTCCATT ATACGAGTCA TGAATATGGT TCAGCATGTT TGGACGCTAC	3900
	TTGTTTGGGG CAATTTTCTG TGGTGAATTG TAGCTGCTTG ATGTTGGCTA	3950
15	GCTGGCTTAT TTTGTACAAG TATCGATGTT AGATGCATAT TTCCTTTTGT	4000
	TCTTGTGCTG TTGCCATGT TGTATTCCCC TTTTCTGTCG CCAGTGTTGC	4050
	ATGTTAAATT GGTTTTTATT ACATAATCAA CTTTGTGCT GACATCAGTC	4100
20	ATTTTTATTG AGCCTTCTTG CTGCAAAATA TAGACCATAC GGTGTTTACA	4150
	GAGATTCCCG CAGCACCTT GTTATACATA ATTTAGCACA TCAGGTTTGG	4200
25	GTCTATCACC TTCATTATC CGTACATGGC TTTGTAAGTC GGTTCACACG	4250
	TATCGTCATA CTGTATGTTA TTTCAATGTC ATTAGGGTGT GGAGCCTGCA	4300
	AGTACATATC CTGATCTGGG ATTGCCACCT GAATGGTATG GAGCTTTAGA	4350
30	ATGGGTATTT CCAGAATGGG CAAGGAGGCA TGCCCTTGAC AAGGGTGAGG	4400
	CAGTTAACTT TTTGAAAGGA GCAGTTGTGA CAGCAGATCG AATTGTGACC	4450
35	GTCAGTCAGG TGAAATACTC AATACTTCTC TTTTTTCTTT GCGGGATGTT	4500
	CTTCAGTTCA ATTGCCCTGT CTTTCACCCA ATTAAGAAAT GATTTAATCT	4550
	TTTGTCTCTA GGGTTATTCA TGGGAGGTCA CAACTGCTGA AGGTGGACAG	4600
40	GGCCTCAATG AGCTCTTAAG CTCCCGAAAA AGTGTATTGA ATGGTAACTA	4650
	TATTTGAATC CACTTATCTT CTTCTGAAAC ATATTACAG AAATAGATGG	4700
45	ATGGGTTGCA AGAATAAATT CAGTTTGCTC TTTCGGTATG AAGGAATTGT	4750
	AAATGGAATT GACATTAATG ATTGGAACCC CACCACAGAC AAGTGTCTCC	4800
	CTCATCATTA TTCTGTCGAT GACCTCTCTG GAAAGGTGTG TGGATAGTAC	4850
50	CCTATATAAT AACATGTATA TCTGATCTAG TACTTTCTTT TTCTTTGCTA	4900
	GTTTGCTTCC CATGATGTTT TCACTAACTA ATCCTATGTG GTTTGGCATA	4950
55	CTTGTCAGGC CAAATGTAAA GCTGAATTGC AGAAGGAGCT GGGTTTACCT	5000



	GTAAGGGAGG ATG TTCCTCT GGTTAGATAC AAACCCCTAA GATATATATT	5050
	TTTTAAATCC CTA AAAAAAAAA CTTGCCGATC ATCTCATTAG CTTGATTAC	5100
5	AGATTGGCTT TATTGGAAGA CTGGATTACC AGAAAGGCAT TGATCTCATT	5150
	AAAATGGCCA TTCCAGAGCT CATGAGGGAG GACGTGCAGT TTGTAAGTTC	5200
	ATATTCTTTT TCTTGAGACT AGAGTATAAA TCAAACATGT AGGTGTGGGG	5250
10	TGGTATAATA CAGACATAAG TTCCAGCTAT TGCTTCCATG AGAATTTTAA	5300
	TGCTATTCAG TAATATGCTA CTGCAAGTTT TGAAACAAAG TTGGAAGCAA	5350
15	TAAATATATG TGTAGCACTG ACCATGCAGT GCCACTATAG CTGGAATGTC	5400
	CTGTAGTCTA TGTGATCTAA CACACTCAAC AACATGTTTT CGCATACAAA	5450
	CACATGCGTG CGCGCAACAA ACATACTCTA CAATAAAATT GGCTTGGTGA	5500
20	ACTGCAGACA TGCTCTTATC TCCATTCCAA CATTCTTGT TTCAACATTG	5550
	GCTGAAGACT AAGAGAAGGG GGACCCAGGG TGATGTAGCC AACTAGATCC	5600
25	AGTAAGGAAG CTAGCCGAGC CTAGGAGGAT TCGCTTAGGT AGCTGGAACG	5650
	TAGGGTCTCT GACAGGGAAG CTTCCGGGAGC TAGTCGATGC AGTGGTGAGG	5700
	AGAGGTGTTG ATATCCTTTG CGTCCAAGAA ACCAAATGTA GGGGACAGAA	5750
30	GGCGAAGGAG GTGGAGGATA CCGGCTTCAA GCTGTGGTAC ATGGGACGGC	5800
	TGCAACAGA AATGGCGTAG GCATCTTGAT CAACAAGAGC CTTAAGTATG	5850
35	GAGTGGTAGA CGTCAAGAGA CGTGGGGACC GGATTATCCT CGTCAAGCTG	5900
	GTAGTTGGGG ACTTAGTTCT CAATGTTATC AGCGTGTATG CCCC GCAAGT	5950
	AGGCCACAAT GAGAACGCCA AGAGGGAGTT CTGGGAAGGC CTGGAAGACA	6000
40	TGGTTAGGAG TGTACCGATT GGCGAGAAGC TCTTCATAGG AGGAGACCTC	6050
	AATGGCCACG TGGGTACATC TAACATAGGT TTTGAAGGGG CACATGGGGG	6100
45	CTTTGGCTAT GGCATCAAGA ATCAAGAAGA AGATGTCTTA CGCTTTGCTC	6150
	TAGCCTACGA CATGATTGTA GCTAACACCC TCTTTAGAAA GAGAGAATCA	6200
	CATCTGGTGA CTTT TAGTAG TGGCCAACAC TAGCCAGATC GATTTCATCC	6250
50	TCTCGAGAAG AGAAGATAGG TGTGCGCGCC TAGACTGCAA GGTGATACCT	6300
	TCGGATTCGT GTCCAGCGGG ATAAGCGTGC CAAAGTCGCT AGAATGAAGT	6350
55	GGTGGAAGCT CAAGGGGGAG GTAGCTCAGG CGTTCAAGGA GAGGGTCATT	6400

	AGGGAGGGCC CTTGGGAGGA AGGAGGGGAT GCGGACAATG TGTGGATGAA	6450
	GATGGCGACT TGCATTCGTA AGGTGGCCTC GGAGGAGTGT GGAGTGTCCA	6500
5	GGGGATGGAG AAGCGAAGAT AAGGATACCT GGTGGTGGA TGATGATGTC	7000
	CAGAAGGCAA TTAAAGAGAA GAAAGATTGC TTTAGACGCC TATACTTGGA	7050
10	TAGGAGTGCA GTCAACATAG AAAAGTACAA GATGGCGAAG AAGGCCGCAA	7100
	AGCGAGCTGT CAGTGAAGCA AGGGGTCGGG CATATGAGGA TCTCTACCAA	7150
	CGGTTAGGCA CGAAGGAAGG CGAAAGGGAC ATCTATAAGA TGGCCAAGAT	7200
	CCGAGAGAGA GGAAGACGAG GGATATTGGC CAAGTCAAAT GCATCAAGGA	7250
15	TGGAGCAGAC CAACTCTTGG TGAAGGACGA GGAGATTAAG CATAGATGGC	7300
	GGGAGTACTT CGACAAGCTG TTCAATGGGG AGGATGAGAG TCCTACCATT	7350
	GAAC TTGACG ACTCCTTTGA TGAGACCATC ATGCGTTTTA TGCGGCGAAT	7400
	CCAGGAGTCC GAGGTCAAGG AGGCTTTAAA AAGGAGGCAA GGCGATGGGC	7450
	CCTGATTGTA TCCCCATTGA GGTGTGGAAA GGCCTCGGGG ACATAGCGAT	7500
20	AGTATGGCTA ACCAAGCTAT TCAACCTCAT TTTTCGGGCA AACAAGATGC	7550
	CAGAAGAATG GAGACGAAGT ATATTAGTAC CAATCATCAA ACAGGGGGGA	7600
	TGTT CAGAGT TGTACTAATT ACCATGGAAT TAAGCTGATG AGCCATACAA	7650
	TGAAGCTATG GGAGAGAATC ATTGAGCACC GCTTAAGAAG AATGACAAGC	7700
	GTGACCAAAA ATCAGTTTGG TTTCATGCCT GGGAGGTCGA CCATGGAAAC	7750
25	CATTTTCTTG GTACGACAAC TTATGGAGAG ATACAGGGAG CAAAAGAAGG	7800
	ACTTG CATAT GGTGTTCAAT GACTTGAAGA AGGCCTATAA TAAGATACCG	7850
	CGGAATGTCA TGTGGTGGGC CTTGGAGAAA CACAAAGTCC CAGCAAAGTA	7900
	CATTACCCTC ATCAAGGACA TGTACGATAA TGTTGTGACA AGTGTTGAA	7950
	CAAGTGATGT CGACACTAAT GACTTCCCGA TTAAGATAGG ACTGCATCAG	8000
30	GGGTCAGCTT TGAGCCCTTA TCTTTTTGCC TTGGTGATGG ATGAGGTCAC	8050
	AAGGGATATA CAAGGAGATA TCCCATGGTG TATGCTCTTT GTGGATGATT	8100
	TGGTGCTAGT TGACGATAGT CGGGCGGGGG TAAATAACAA GTTAGAGTTA	8150
	TGGAGACAAA CCTTGGAATC GAAAGGGTTT AGGCTTAGTA GAACTAAAC	8200
	CGAGTACATG ATGTGCGGTT TCAGTACTAC TAGGTGTGAG GAGGAGGAGG	8250

- 93 -

	TTAGCCTTGA TGGGCAGGTG GTACCCCAAG AGGACACCTT TCGATATTTG	8300
	GGGTCAATGC TGCAGGAGGA TGGGGGTATT GATGAAGATG TGAACCATCG	8350
	AATCAAAGCT GGATGGATGA AGTGGCGCCA AGCTTCTGGC ATTCTTTGTG	8400
	ACAAGAGAGT GCCACAAAAG CTAAGGCAAG TTCTACAGGA CGGCGGTTCCG	8450
5	ACCCGCAATG TTGTATGGCG CTGAGTGTG GCCGACTAAA AGGCGACATG	8500
	TTCAACAGTT AGGTGTGGCG GAGATGCGTA TGTTGAGATG GATGTGTGGC	8550
	CACACGAGGA AGRATCGAGT CCGGAATGAT GATATACGAG ATAGAGTTGG	8600
	GGTAGCACCA ATTGAAGAGA AGCTTGTCCA ACATCGTCTG AGATGGTTTG	8650
	GGCATATTCA GCGCACGCCT CCGAAAACCT CAGTGCATAA CGGACGGCTA	8700
10	AAGCGTGCGG AGAATGTCAA GAGAGGGCGG GGTAGACCGA ATTTGACATG	8750
	GGAGGAGTCC GTTAAGAGAG ACCTGAAGGT TTGGAGTATT ACGAAAGAAC	8800
	TAGCTATGGA CARGGGTGCG TGGAAGCTTG TTATCCATGT GCCAGAGCCA	8850
	TGAGTTGATC ACGAGATCTT ATGGGTTTCA CCTCTAGCCT ACCCCAACCT	8900
	GTTTGGGACT AAAGGCTTTG TTGTTGTTGT TGTTGTTGTT GTTGTAGCCA	8950
15	ACTAAATCCA GTTGATCAGT GGTTTTTACT CTTATTTTTA CAGGTCATGC	9000
	TTGGATCTGG GGATCCAATT TTTGAAGGCT GGATGAGATC TACCGAGTCG	9050
	AGTTACAAGG ATAAATTCCG TGGATGGGTT GGATTTAGTG TTCCAGTTTC	9100
	CCACAGAATA ACTGCAGGGT ATGCCGAGAA CTTCTTAACA AGACCTTCGT	9150
	TATCAGCTTG GATATATTAT AATGTTCAAA ACATTTATGT CTCTCTTTTT	9200
20	GTGCAGTTGC GATATATTGT TAATGCCATC CAGGTTTGAA CCTTGTGGTC	9250
	TTAATCAGCT ATATGCTATG CAATATGGTA CAGTTCCTGT AGTTCATGGA	9300
	ACTGGGGGCC TCCGAGTAAG ACAACTGCCT TGAAAATTAT CGTTATCTTG	9350
	GCTCCAACGC AAATGTTCTA ATTGGCTCGT GTATTCAACA GGACACAGTC	9400
	GAGACCTTCA ACCCTTTTGG TGCAAAAGGA GAGGAGGGTA CAGGGTACGC	9450
25	ACTGCTCAAT TTTAGCTAAC TTTCAGTTTA TCTTTTGTCA ATGTCTTGGG	9500
	GGTTCATTGC GCCATAAATC AACTTGTGAT AATTAAGTGT TACTGTTCTG	9550
	TACTTGCAGG TGGGCGTTCT CACCGCTAAC CGTGGACAAG ATGTTGTGGG	9600
	TAAGTTTTTG CTGAGCTCTT GTCCGGTTAT AGGATCGACC TTGGCTGTAG	9650

- 94 -

	CATGGTACCT TAGTGCCCT TGTATATAGA CCTAACCTGA TGGACTCACT	9700
	TTGTCTACAC TAATCATAGT AGTCGATTGC CCGGAGGCGT TTTGCTTGGA	9750
	TTCTGCTAAT TTAATTTTCA TGACGATAAC TCATACCATG GTTTGGTTCT	9800
	CCGATGGGGG CCAGAATGGC GTCTAGTGTC TGCGATCTGT GTAAC TAGCC	9850
5	AATGCCGGGT TGTTCCAAGT GAAAATTTAC CTTTGGACCA TTGTGCAGGC	9900
	ATTGCGAACC GCGATGTCGA CATTGAGGGA GCACAAGCCG TCCTGGGAGG	9950
	GGCTCATGAA GCGAGGCATG ACGAAAGACC ATACGTGGGA CCA <sup>1</sup> TGCCGCC	10000
	GAGCAGTACG AGCAGATCTT CGAATGGGCC TTCGTGGACC AACCTACGT	10050
	CATGTAGACG GGGACTGGGG AGGTCGAAGC GCGGGTCTCC TTGAGCTCTG	10100
10	AAGACATGTT CCTCATCCTT CCGCGGCCCG GAAGGATACC CCTGTACATT	10150
	GCGTTGTCCT GCTACAGTAG AGTCGCAATG CGCCTGCTTG CTTGGTCCGC	10200
	CGGTTGCGA G TAGATGACG GCTGTGCTGC TGCGGCGGTG ACAGCTTCGG	10250
	GTGGATGACA GTTACAGTTT TGGGGAATAA GGAAGGGATG TGCTGCAGGA	10300
	TGGTTAACAG CAAAGCACCA CTCAGATGGC AGCCTCTCTG TCCGTGTTAC	10350
15	AGCTGAAATC AGAAACCAAC TGGTGA <sup>1</sup> CTCT TTAGCCTTAG CGATTGTGAA	10400
	GTTTGTTGCA TTCTGTGTAT GTTGTCTTGT CCTTAGCTGA CAAATATTTG	10450
	ACCTGTTGGA TAATTCTATC TTTGCTGCTG TTTTCTTTT GGTCAAAAGA	10500
	GGGGTTCCT CCGATTTTCA <sup>1</sup> T TAACGAAACC ACCAAAATAA CAGCACCCAG	10550
	TGCAGGTCTC AGGTTGAGAT ATACTTAAGA CTA <sup>1</sup> CTAAATC TAACAGCAGC	10600
20	TAAAAAGCTT AAAGATTCAG GCGACATAAC CGAACAAAAT CCACAACCGA	10650
	AGGGACCAAA GCAGGACAAG TAAAAAGGCA GNCGACACAA AGCGCAGGTC	10700
	GCTGAAAAGG CAAGCAGACA GAGGTCTGCA TTCTGTCAAC ACCACTTGTC	10750
	AAAAATGAAG AGAAGATCGA GAATTC <sup>1</sup> CCCGG GAATCCG	10787

25 (2) INFORMATION FOR SEQ ID NO: 14:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 647 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

30

(ii) MOLECULE TYPE: protein

(iii) HYPOTHETICAL: NO

- 95 -

## (vi) ORIGINAL SOURCE:

(A) ORGANISM: triticum tauschii

(F) TISSUE TYPE: Endosperm

## 5 (ix) FEATURE:

(A) NAME/KEY: Protein

(B) LOCATION: 1..647

(D) OTHER INFORMATION: /product= "deduced amino acid  
sequence for SSS I"

10

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 14:

```

Met Ala Ala Thr Gly Val Gly Ala Gly Cys Leu Ala Pro Ser Val Arg
1          5          10          15
Leu Arg Ala Asp Pro Ala Thr Ala Ala Arg Ala Ser Ala Cys Val Val
20          25          30
Arg Ala Arg Leu Arg Arg Leu Ala Arg Gly Arg Tyr Val Ala Glu Leu
35          40          45
Ser Arg Glu Gly Pro Ala Ala Arg Pro Ala Gln Gln Gln Gln Leu Ala
50          55          60
Pro Pro Leu Val Pro Gly Phe Leu Ala Pro Pro Pro Pro Ala Pro Ala
65          70          75          80
Gln Ser Pro Ala Pro Thr Gln Pro Pro Leu Pro Asp Ala Gly Val Gly
85          90          95
Glu Leu Ala Pro Asp Leu Leu Leu Glu Gly Ile Ala Glu Asp Ser Ile
100          105          110
Asp Ser Ile Ile Val Ala Ala Ser Glu Gln Asp Ser Glu Ile Met Asp
115          120          125
Ala Asn Glu Gln Pro Gln Ala Lys Val Thr Arg Ser Ile Val Phe Val
130          135          140
Thr Gly Glu Ala Ala Pro Tyr Ala Lys Ser Gly Gly Leu Gly Asp Val
145          150          155          160
Cys Gly Ser Leu Pro Ile Ala Leu Ala Ala Arg Gly His Arg Val Met
165          170          175
Val Val Met Pro Arg Tyr Leu Asn Gly Ser Ser Asp Lys Asn Tyr Ala
180          185          190
Lys Ala Leu Tyr Thr Gly Lys His Ile Lys Ile Pro Cys Phe Gly Gly
195          200          205
Ser His Glu Val Thr Phe Phe His Glu Tyr Arg Asp Asn Val Asp Trp
210          215          220
Val Phe Val Asp His Pro Ser Tyr His Arg Pro Gly Ser Leu Tyr Gly
225          230          235          240
Asp Asn Phe Gly Ala Phe Gly Asp Asn Gln Phe Arg Tyr Thr Leu Leu
245          250          255
Cys Tyr Ala Ala Cys Glu Ala Pro Leu Ile Leu Glu Leu Gly Gly Tyr
260          265          270

```

- 96 -

	Ile Tyr Gly Gln Asn Cys Met Phe Val Val Asn Asp Trp His Ala Ser	275	280	285
5	Leu Val Pro Val Leu Leu Ala Ala Lys Tyr Arg Pro Tyr Gly Val Tyr	290	295	300
10	Arg Asp Ser Arg Ser Thr Leu Val Ile His Asn Leu Ala His Gln Gly	305	310	315
	Leu Glu Pro Ala Ser Thr Tyr Pro Asp Leu Gly Leu Pro Pro Glu Trp	325	330	335
15	Tyr Gly Ala Leu Glu Trp Val Phe Pro Glu Trp Ala Arg Arg His Ala	340	345	350
	Leu Asp Lys Gly Glu Ala Val Asn Phe Leu Lys Gly Ala Val Val Thr	355	360	365
20	Ala Asp Arg Ile Val Thr Val Ser Gln Gly Tyr Ser Trp Glu Val Thr	370	375	380
25	Thr Ala Glu Gly Gly Gln Gly Leu Asn Glu Leu Leu Ser Ser Arg Lys	385	390	395
	Ser Val Leu Asn Gly Ile Val Asn Gly Ile Asp Ile Asn Asp Trp Asn	405	410	415
30	Pro Thr Thr Asp Lys Cys Leu Pro His His Tyr Ser Val Asp Asp Leu	420	425	430
	Ser Gly Lys Ala Lys Cys Lys Ala Glu Leu Gln Lys Glu Leu Gly Leu	435	440	445
35	Pro Val Arg Glu Asp Val Pro Leu Ile Gly Phe Ile Gly Arg Leu Asp	450	455	460
40	Tyr Gln Lys Gly Ile Asp Leu Ile Lys Met Ala Ile Pro Glu Leu Met	465	470	475
	Arg Glu Asp Val Gln Phe Val Met Leu Gly Ser Gly Asp Pro Ile Phe	485	490	495
45	Glu Gly Trp Met Arg Ser Thr Glu Ser Ser Tyr Lys Asp Lys Phe Arg	500	505	510
	Gly Trp Val Gly Phe Ser Val Pro Val Ser His Arg Ile Thr Ala Gly	515	520	525
50	Cys Asp Ile Leu Leu Met Pro Ser Arg Phe Glu Pro Cys Gly Leu Asn	530	535	540
55	Gln Leu Tyr Ala Met Gln Tyr Gly Thr Val Pro Val Val His Gly Thr	545	550	555
	Gly Gly Leu Arg Asp Thr Val Glu Thr Phe Asn Pro Phe Gly Ala Lys	565	570	575
60	Gly Glu Glu Gly Thr Gly Trp Ala Phe Ser Pro Leu Thr Val Asp Lys	580	585	590
	Met Leu Trp Ala Leu Arg Thr Ala Met Ser Thr Phe Arg Glu His Lys	595	600	605

- 97 -

Pro Ser Trp Glu Gly Leu Met Lys Arg Gly Met Thr Lys Asp His Thr  
 610 615 620

5 Trp Asp His Ala Ala Glu Gln Tyr Glu Gln Ile Phe Glu Trp Ala Phe  
 625 630 635 640

Val Asp Gln Pro Tyr Val Met  
 645

10

## (2) INFORMATION FOR SEQ ID NO: 15:

## (i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 5072 base pairs

(B) TYPE: nucleic acid

15

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

20

(iii) HYPOTHETICAL: NO

(vi) ORIGINAL SOURCE:

(A) ORGANISM: triticum tauschii

(F) TISSUE TYPE: Endosperm

25

(ix) FEATURE:

(A) NAME/KEY: promoter

(B) LOCATION: 1..4993

(D) OTHER INFORMATION: /function= "region containing promoter of SSS I"

30

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 15:

TCTAGATGCA TGCTGGATAG CGGTCGATGT GTGGAGTAAT AGTAGTAGAT GCAGAATCGT 60

35 TTCGGTCTAC TTGTCGCGGA CGTGATGCCT ATATACATGA TCATACCTAG ATATTCTCAT 120

AACTATGCTC AATTCTATCA ATTGCTCGAC AGTAATTCGT TTACCCACCG TAATACTTAT 180

40 GATCTTGAGA GAAGTCACTA GTGAAACCTA TGCCCCCAG GTCTATTTTG CATCATATTA 240

ATCTTCCAAT ACTTAGTTAT TTCCATTGCC GTTTATTTTA CTTTGTATCT TTATTTCTTT 300

TTATTATAAA AAATACCAA AATATTATCT TATCATATCT ATCAGATCTC ATTCTCGTAA 360

45 GTGACCGTGA AGGGATTGAC AACCCCTTTA TCGTGTTGGT TGCGAGGTTT TTGTTTGTTT 420

GTGTAGGTGC GTGTGACTCG CACGTCTCCT ACTGGATTGA TACCTTGGGT TTTCAAAAAC 480

50 TGAGAAAAAT ACTTACGCTA CTTTACTGCA TAACCCTTTC CTCTTTAAAA AAAAAAACCA 540

ACGTAGTATT CAAGAGGTAG CACGCTACCA TCCTCTCCAA CAGGAGCGCG GAGATCTTTG 600

TCCGGCAGGT TGATGCGGGC CGGGGAAGAA CTCCAGCTGC CTTGGCCAGC TTGGTCGTGA 660

55 GCCGCCCCAG CGGCGTCTTG AACCTGTCCA CGTAGCGCTC CCTGACACGC GCGTGAAC 720

GAGAAGGCTT GTCGATGAAC TCCAGCTGTT GTGCCAGCCT AGCTTGCGCC TTCTTCTGCT 780

60 GGGTCATGCC CTTCGAGAAA CCCACCTTGG CCACCCTTGT GCTTGAGCGG CGCGCCACCT 840

CAGCAGGCGG CGGCGTGGGG ATGAAGAGGG TGTCTGCTTC CGGAGCAGGC GGGTCGGCGT 900

- 98 -

TGAACTTGAA AGGCGGTGGC CCCATGATGG ATGGGGGGAG CATGCCAAAG ACTTGGTTGA 960  
GGAAAGTGGT GTTGGCGTCC ACCTCCAGTG CCTGCAGTTT GGAAGCCAGA CGATTGGCGT 1020  
5 CGATCTCTGG CTCCGGCTGG AAGGAGGCTC GACGCTCCGG TGTGCCAGAA CGCAAAGGGA 1080  
GGAGCGGCAG CTCTGGCTGA GCAGACCCCG CGCCCATGTA CTCTGCATTG GGCCAAGGCT 1140  
GCAGGGGCAA GCCACCGGGA TGGGGGCGCG AGGTGGACTG CGCACCGGAG GAAGGCCAAG 1200  
10 CTCAACCTCG GTGAGGTTTCG CCCCAGACCA GGGCGGCAGG CTCGGGTCCA CAAAGGGCCA 1260  
AACCGCCTCG TCCGCCCGA AACTGTCCAG GACAGACGGC GGACGACGGA AGGCCGTGTC 1320  
15 GTCGAGCTCG AGCAGCAGAG GGTCCGTGCG GGTGATGTCT TGCCAAATGG ACTCCACCTC 1380  
CAGCAGGAAG GGGGACTGGT CCATCGCCCC TGGCCAAGCC ACTGGTACGC CAAAGATGGC 1440  
ATCAGCAGCG TTTGCACCAG GGGGAGCAGC CACACCTTGG AGGACAGGGA GGGTGCGGAC 1500  
20 GTCGACGGCA GCAAAACGTG GCTGGAGCAA GTTGCCGTCG CGTGCCGGCC TCGGCGAGCG 1560  
CGAGCGGCTG TAGGAGCGCT CGGTGCCCTC AGACTCGGAC AGTGCGCCAG TGGGAGAGCC 1620  
25 ATGGCGACGC CGGCCACCAC TGGACGTGCC ATGGCGCTGG TCCTGACGGC GCCTGGATGG 1680  
CCCGTCCTCG CGGGCAGCTC CACCTGAGCG GCACCCGAGG AGCACACCCC GCCAAGCTGG 1740  
GCCAGGGCGG CTGCGGCGAC GGCACGGCC GCGGTCGCGG TCTGCACCAT CATCTTCATC 1800  
30 TTCGTATCG TGGCGCCTCG GACAAGGATG CTCGCTGTCA CCGACGCGAG GGACGTGAGC 1860  
CGGCTCAGCC CGCCCTTCCT CGACGTGGCG AGCCCTGCGG ATATGCTCCT CGAGCGGCCA 1920  
35 TTGGGGGTCG TTGGCGCGCG GCATCTCGGG GTCGCGGTCA GCTATCGGGG TGTAGTCCTT 1980  
TGTGGTGTCC AGGTGGATGA GCAGAGAGAA ATCCGGCCCC TCTAGCCCCCT CGTCCCGGGG 2040  
GCAGCCCTCC GGCAGCGTCT GCGGCCCCCT GGGGTCCAGG GGTGATCGA TGATGGAGAA 2100  
40 CCCCCTTTTG GTGGGGATGT CGTCCGGA CTATGCCAC ACCCAGGCAA AGAGGCAGGC 2160  
CGTGTGGAG AGGGAGGTCG TCTGCCGCTC CAACAGTCG ACGTGGCATG TCTTCCCGAG 2220  
45 CGCATCTGCG CCCGCTCCT TGTTCAGGA CTGCACCGGC ATGTTCTCGA CGGCGATGCG 2280  
GCAGTAGTAC CGCCAGACAC GCGGTGGCC GTGTGCCGAT GGTGACCAGG CCGACAGGGA 2340  
GAGCGCGACG CCCCAGCAGG AGACGACCCC AGCGTCGAAA GCGATGTCCC GGTGCCTGAA 2400  
50 GTGGACGAGC CCAGAGATGG CCAGGCGCAT TGACGCGGGG AAGGGGAAGG AGTTAGGATG 2460  
GGCGACGCGG CCGGAGTGAA CCGCGGCGTG GTGGCCGACG GGGCTGGAGA GGCAGAGGCG 2520  
55 GAGTCATCCG AGAGAGGTGT ATCAGTGGCT CTGCACAATA CCCAGTGTGCG CCACATCATA 2580  
TCCTGCTGAA TAACCACACA TGTGTACTGT CGTTAAATAA ATCATTGGTC ACGCGAACCC 2640  
GGAAAAAGAC GGCGAAAAAT TCACGGACAC ACGACTAGTA GTACCCAATA TACTCGGCAA 2700  
60 AAACAGTGAC ACGTCGTTTT GCGTTGTCGG CCGGTGTTGT CGAGTCATTG TACTATGTTT 2760  
TGTCGTTTCT TTCTTTTCTC CAAATCGACA AACC GTTTGT CTTTGGTTAA AAAACAGAAA 2820  
65 CATACAAAAT CAAATGAATG CATTCAAGGG CCGTAATCC AATTCTGAGC CCAGGCTCAG 2880  
CTACACCCGC CCTTACAAAA AAATCAAAAT AAATACTAGA AAAATTCAAA AAATTCCAAT 2940



TTGTTTGTGC GTGGTAGATA ATTTGATGCG TGAGGTACGC TTCAATTTTC AAATTATTTG 3000  
5 GACATCTGAG CAGCTCTCAG CAAAAAAGAC AAATTCGGGG TCTGTAAAAA TGTTTACTGT 3060  
TCATGCACTG TTCTGACCCG ATTTGTCTTT TTTGCTGAGA GCTTCTCAGA AGTCCAAATG 3120  
AGCTAAAATT TTGAGCGGAG CTTACGTGAT AAAATGTCTA TCATGCAAAA AAGGATTGGA 3180  
10 ATTTTTTGAA TTTTTTTTAT TTTTTGTGAT TTGTTTCCTG GACGGGTGCA GATAAGCCTG 3240  
GGCACCGAAA CGCCGCACTC AGGCTCATCC TTTTCTATAA AAGAAAAGAA ATACATACAA 3300  
15 TTTCCCTCTG TTTTTTGAGC AAGGGGCACC ACCCACCAAA GAGTTTTCOA CTCACATGGT 3360  
ATTAGAGCAT CTACAGCCGG GCGTCTCAAA CCAGCCTCAT ACGCTTGAGC GGGTCGCCTT 3420  
GGTCACGATT TTTTGACCCA GACGGGCCCC TCAAACGGTC CTAAACGCC CAGGCTGACC 3480  
20 GACAACCCAC ATATCCAGCC CAAATATGGG GTGGATATGG GGGCGCCCGG GCACGCCAGC 3540  
CCGCGGACAC CACACATCTT CAGTTTCTAA TTTGAGATAT CCGGATGTGG AATGCGTTTT 3600  
25 TGAGGGGTGA CCGGTCCCTG TCCGTGGATG CGCCCGGACG TTTGAGGGGT TGGATTTGCC 3660  
AAGTCTGATT AGAGATGCTC TTAGGTGTTT CACCCCCATC CCTTGATGGC TAGGGCAAAC 3720  
TCTCCCTCC AAACCTTGTC GCGGAGCCTG TGGATTCTTC TCTCCTCTGC CCGCTGCTCC 3780  
30 GCGGCTGAT GCGGGGAGG AGAATCCCG TGTCTTCGCT TGGTTAGTTG TTTAAGTTAC 3840  
GTACTTTTTT AGTCCTCGCA GGTGCGGCGT TCGGACGTAT GGTCGTGCTT CTTTTTTGAG 3900  
35 TTTGTCTTCC GGGCTCTGAT CCTCCTCGAG TTCGTCCATC TGGACGTACT CGACGGAGCT 3960  
CCGGCATAGA TTCCTATCAT CGTCTTGGTG AGGTGAGGTT ATGGTTTCTT GTCATGTGGG 4020  
CAGATTTGGT GCCAGATGCT TCATATCTAT TCAAGGGTTC AGCGGCAACA ACTGCGGCTC 4080  
40 CAGAGCGATG GTCCTTAAGG GCACGTGCAC GAAGACTTCA CGGCTGTTAT CGACAAGGTC 4140  
AAGCCGGCTC CGATAGGGGA GCAGCGACAG CGGCGCGTCA ACCGCTCGTT CTGGCGGCAG 4200  
45 TAGTGGTCTG TCGGTGCTCT CGGAACCTCG ATGTAATTTT TATGATTTTA GAGATGCTTT 4260  
GTACTTCCGA TCGATGAACT CTGATAATAG ATATCTCTTC TCTCGCAAAA AAAGAGAGTT 4320  
TTCAACTGAA AACAAAAGAG TTTCACTAGT TCTTCTTTTA GAAACAGAGT TTCCTAGCA 4380  
50 CTTTTTTTTG CGAGAAGTCG AGTTTCACTA AGTACTAAAC CCACGCAATT ATTCTCAAAA 4440  
AAAAAACCCA CGCAACTGTC TGGATCCATC TTCGTTTTTT CCCCAGAGAT CGTCTGGATC 4500  
55 CATTTTCGTG TCGGAGGCAT CCTCTCATTT TGCACGGCCC AGCTCTCTTC TCGCCGGCGT 4560  
ACGCTGCTAC ATGTCGGCAC TCCACGCAA CAAAAAGAAG CCCAACCGAA AACGCACGCG 4620  
CCTTTCCAGG CTCACCACGG AAAAAAATAC CACGCGCCGC TCACGAGCAA ACCGTGACAA 4680  
60 CAGCCAGCCA GATATGGCAA CGGAGGCACG GGCCGCACAC AGCCACTGAA AACCGCAGCT 4740  
GCTCTTCCGT CCGTCCGTCC CTCCGCCCGT CCGCGCCACT CCACTCGCCT TGCCCCACTC 4800  
65 CCACTCTTCT CTCCCCGCGC ACACCGAGTC GGCACCGGCT CATCACCCAT CACCTCGGCC 4860  
TCGGCCACCG GCAAACCCCC CGATCCGCTT TTGCAGGCAG CGCACTAAAA CCCCAGGGAG 4920

- 100 -

CGCGCCCCGC GGCAGCAGCA GCACCGCAGT GGGAGAGAGA GGCTTCGCCC CGGCCCCGCAC 4980  
 CGAGCGGGGC GATCCACCGT CCGTGCGTCC GCACCTCCTC CGCCTCCTCC CCTGTCCCCG 5040

5 GCGCCACAC CCATGGCGGC GACGGGCGTC GG 5072

(2) INFORMATION FOR SEQ ID NO: 16:  
 (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 1706 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(vi) ORIGINAL SOURCE:  
 (A) ORGANISM: *triticum tauschii*  
 (F) TISSUE TYPE: Endosperm

(ix) FEATURE:  
 (A) NAME/KEY: CDS  
 (B) LOCATION: 1..1706  
 (D) OTHER INFORMATION: /product= "partial cDNA for hexaploid wheat DBE"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 16:

30	GCT GTG TCG AAG CTT GAC TAT TTG AAG GAG CTT GGA GTT AAT TGT ATT	48
	Ala Val Ser Lys Leu Asp Tyr Leu Lys Glu Leu Gly Val Asn Cys Ile	
	1 5 10 15	
35	GAA TTA ATG CCC TGC CAT GAG TTC AAC GAG CTG GAG TAC TCA ACC TCT	96
	Glu Leu Met Pro Cys His Glu Phe Asn Glu Leu Glu Tyr Ser Thr Ser	
	20 25 30	
40	TCT TCC AAG ATG AAC TTT TGG GGA TAT TCT ACC ATA AAC TTC TTT TCA	144
	Ser Ser Lys Met Asn Phe Trp Gly Tyr Ser Thr Ile Asn Phe Phe Ser	
	35 40 45	
45	CCA ATG ACG AGA TAC ACA TCA GGC GGG ATA AAA AAC TGT GGG CGT GAT	192
	Pro Met Thr Arg Tyr Thr Ser Gly Gly Ile Lys Asn Cys Gly Arg Asp	
	50 55 60	
50	GCC ATA AAT GAG TTC AAA ACT TTT GTA AGA GAG GCT CAC AAA CGG GGA	240
	Ala Ile Asn Glu Phe Lys Thr Phe Val Arg Glu Ala His Lys Arg Gly	
	65 70 75 80	
55	ATT GAG GTG ATC CTG GAT GTT GTC TTC AAC CAT ACA GCT GAG GGT AAT	288
	Ile Glu Val Ile Leu Asp Val Val Phe Asn His Thr Ala Glu Gly Asn	
	85 90 95	
60	GAG AAT GGT CCA ATA TTA TCA TTT AGG GGG GTC GAT AAT ACT ACA TAC	336
	Glu Asn Gly Pro Ile Leu Ser Phe Arg Gly Val Asp Asn Thr Thr Tyr	
	100 105 110	
60	TAT ATG CTT GCA CCC AAG GGA GAG TTT TAT AAC TAT TCT GGC TGT GGG	384
	Tyr Met Leu Ala Pro Lys Gly Glu Phe Tyr Asn Tyr Ser Gly Cys Gly	
	115 120 125	

- 101 -

	AAT	ACC	TTC	AAC	TGT	AAT	CAT	CCT	GTG	GTT	CGT	CAA	TTC	ATT	GTA	GAT	432
	Asn	Thr	Phe	Asn	Cys	Asn	His	Pro	Val	Val	Arg	Gln	Phe	Ile	Val	Asp	
	130						135					140					
5	TGT	TTA	AGA	TAC	TGG	GTG	ATG	GAA	ATG	CAT	GTT	GAT	GGT	TTT	CGT	TTT	480
	Cys	Leu	Arg	Tyr	Trp	Val	Met	Glu	Met	His	Val	Asp	Gly	Phe	Arg	Phe	
	145					150					155					160	
10	GAT	CTT	GCA	TCC	ATA	ATG	ACC	AGA	GGT	TCC	AGT	CTG	TGG	GAT	CCA	GTT	528
	Asp	Leu	Ala	Ser	Ile	Met	Thr	Arg	Gly	Ser	Ser	Leu	Trp	Asp	Pro	Val	
					165					170					175		
15	AAC	GTG	TAT	GGA	GCT	CCA	ATA	GAA	GGT	GAC	ATG	ATC	ACA	ACA	GGG	ACA	576
	Asn	Val	Tyr	Gly	Ala	Pro	Ile	Glu	Gly	Asp	Met	Ile	Thr	Thr	Gly	Thr	
				180						185					190		
20	CCT	CTT	GTT	ACT	CCA	CCA	CTT	ATT	GAC	ATG	ATC	AGC	AAT	GAC	CCA	ATT	624
	Pro	Leu	Val	Thr	Pro	Pro	Leu	Ile	Asp	Met	Ile	Ser	Asn	Asp	Pro	Ile	
				195				200					205				
25	CTT	GGA	GGC	GTC	AAG	CTC	ATT	GCT	GAA	GCA	TGG	GAT	GCA	GGA	GGC	CTC	672
	Leu	Gly	Gly	Val	Lys	Leu	Ile	Ala	Glu	Ala	Trp	Asp	Ala	Gly	Gly	Leu	
	210						215					220					
30	TAT	CAA	GTA	GGT	CAA	TTC	CCT	CAC	TGG	AAT	GTT	TGG	TCT	GAG	TGG	AAT	720
	Tyr	Gln	Val	Gly	Gln	Phe	Pro	His	Trp	Asn	Val	Trp	Ser	Glu	Trp	Asn	
	225					230					235					240	
35	GGG	AAG	TAC	CGG	GAC	ATT	GTG	CGC	CAA	TTC	ATT	AAA	GGC	ACT	GAT	GGA	768
	Gly	Lys	Tyr	Arg	Asp	Ile	Val	Arg	Gln	Phe	Ile	Lys	Gly	Thr	Asp	Gly	
					245					250					255		
40	TTT	GCT	GGT	GGT	TTT	GCC	GAA	TGT	CTT	TGT	GGA	AGT	CCA	CAC	CTA	TAC	816
	Phe	Ala	Gly	Gly	Phe	Ala	Glu	Cys	Leu	Cys	Gly	Ser	Pro	His	Leu	Tyr	
				260					265						270		
45	CAG	GCA	GGA	GGA	AGG	AAA	CCT	TGG	CAC	AGT	ATC	AAC	TTT	GTA	TGT	GCA	864
	Gln	Ala	Gly	Gly	Arg	Lys	Pro	Trp	His	Ser	Ile	Asn	Phe	Val	Cys	Ala	
				275				280					285				
50	CAT	GAT	GGA	TTT	ACA	CTG	GGT	GAT	TTG	GTA	ACA	TAT	AAT	AAC	AAG	TAC	912
	His	Asp	Gly	Phe	Thr	Leu	Gly	Asp	Leu	Val	Thr	Tyr	Asn	Asn	Lys	Tyr	
	290						295					300					
55	AAT	TTA	CCA	AAT	GGG	GAG	AAC	AAT	AGA	GAT	GGA	GAA	AAT	CAC	AAT	CTT	960
	Asn	Leu	Pro	Asn	Gly	Glu	Asn	Asn	Arg	Asp	Gly	Glu	Asn	His	Asn	Leu	
	305					310					315					320	
60	AGC	TGG	AAT	TGT	GGG	GAG	GAA	GGA	GAA	TTC	GCA	AGA	TTG	TCT	GTC	AAA	1008
	Ser	Trp	Asn	Cys	Gly	Glu	Glu	Gly	Glu	Phe	Ala	Arg	Leu	Ser	Val	Lys	
					325					330					335		
65	AGA	TTG	AGG	AAG	AGG	CAG	ATG	CGC	AAT	TTC	TTT	GTT	TGT	CTC	ATG	GTT	1056
	Arg	Leu	Arg	Lys	Arg	Gln	Met	Arg	Asn	Phe	Phe	Val	Cys	Leu	Met	Val	
				340					345						350		
70	TCT	CAA	GGA	GTT	CCA	ATG	TTT	TAC	ATG	GGC	GAT	GAA	TAT	GGC	CAC	ACA	1104
	Ser	Gln	Gly	Val	Pro	Met	Phe	Tyr	Met	Gly	Asp	Glu	Tyr	Gly	His	Thr	
				355				360					365				
75	AAA	GGG	GGC	AAC	AAC	AAT	ACA	TAC	TGC	CAT	GAT	TCT	TAT	GTC	AAT	TAT	1152
	Lys	Gly	Gly	Asn	Asn	Asn	Thr	Tyr	Cys	His	Asp	Ser	Tyr	Val	Asn	Tyr	
	370						375					380					

- 102 -

5	TTT CGC TGG GAT AAA AAA GAA CAA TAC TCT GAC TTG CAC AGA TTC TGC 1200 Phe Arg Trp Asp Lys Lys Glu Gln Tyr Ser Asp Leu His Arg Phe Cys 385 390 395 400
10	TGC CTC ATG ACC AAA TTC CGC AAG GAG TGC GAG GGT CTT GGC CTT GAG 1248 Cys Leu Met Thr Lys Phe Arg Lys Glu Cys Glu Gly Leu Gly Leu Glu 405 410 415
15	GAC TTT CCA ACG GCC GAA CGG CTG CAG TGG CAT GGT CAT CAG CCT GGG 1296 Asp Phe Pro Thr Ala Glu Arg Leu Gln Trp His Gly His Gln Pro Gly 420 425 430
20	AAG CCT GAT TGG TCT GAG AAT AGC CGA TTC GTT GCC TTT TCC ATG AAA 1344 Lys Pro Asp Trp Ser Glu Asn Ser Arg Phe Val Ala Phe Ser Met Lys 435 440 445
25	GAT GAA AGA CAG GGC GAG ATC TAT GTG GCC TTC AAC ACC AGC CAC TTA 1392 Asp Glu Arg Gln Gly Glu Ile Tyr Val Ala Phe Asn Thr Ser His Leu 450 455 460
30	CCG GCC GTT GTT GAG CTC CCA GAG CGC GCA GGG CGC CGG TGG GAA CCG 1440 Pro Ala Val Val Glu Leu Pro Glu Arg Ala Gly Arg Arg Trp Glu Pro 465 470 475 480
35	GTG GTG GAC ACA GGC AAG CCA GCA CCA TAT GAC TTC CTC ACC GAC GAC 1488 Val Val Asp Thr Gly Lys Pro Ala Pro Tyr Asp Phe Leu Thr Asp Asp 485 490 495
40	TTA CCT GAT CGC GCT CTC ACC ATA CAC CAG TTC TCT CAT TTC CTC AAC 1536 Leu Pro Asp Arg Ala Leu Thr Ile His Gln Phe Ser His Phe Leu Asn 500 505 510
45	TCC AAC CTC TAC CCC ATG CTC AGC TAC TCA TCG GTC ATC CTA GTA TTG 1584 Ser Asn Leu Tyr Pro Met Leu Ser Tyr Ser Ser Val Ile Leu Val Leu 515 520 525
50	CGC CCT GAT GTT TGA GAG ACA AAT ATA TAC AGT AAA TAA TAT GTC TAT 1632 Arg Pro Asp Val * Glu Thr Asn Ile Tyr Ser Lys * Tyr Val Tyr 530 535 540
55	ATG TAG TCC TTT GGC GTA TTA TCA GTG TGC ACA ATT GCT CTA TTG CCA 1680 Met * Ser Phe Gly Val Leu Ser Val Cys Thr Ile Ala Leu Leu Pro 545 550 555 560
60	GTG ATC TAT TCG ATA GCG GCC GCG AA 1706 Val Ile Tyr Ser Ile Ala Ala Ala 565

## 50 (2) INFORMATION FOR SEQ ID NO: 17:

## (i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 9289 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

55 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

60

(vi) ORIGINAL SOURCE:

- 103 -

(A) ORGANISM: triticum tauschii

(F) TISSUE TYPE: Endosperm

(ix) FEATURE:

5 (A) NAME/KEY: CDS

(B) LOCATION:1..9289

(D) OTHER INFORMATION:/product= "genomic sequence of DBE"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 17:

10	CGG GAC CGT CCC TTG GCA ACT TGG GTT ACG TTG GGA CCT GAC GCT TCG	48
	Arg Asp Arg Pro Leu Ala Thr Trp Val Thr Leu Gly Pro Asp Ala Ser	
	570 575 580	
15	CTT ATC CGG TGT GCC CTG AGA CGA GAT ATG TGC AGC TCC TAT CGG ATT	96
	Leu Ile Arg Cys Ala Leu Arg Arg Asp Met Cys Ser Ser Tyr Arg Ile	
	585 590 595 600	
20	TGT CGG CAC ATT CGG CGG CTT TGC TGG TCT TGT TTT ACC ATT GTC GAA	144
	Cys Arg His Ile Arg Arg Leu Cys Trp Ser Cys Phe Thr Ile Val Glu	
	605 610 615	
25	ATG TCT TAT AAA CCG GGA TTC CGA GAC TGA TCG GGT CTT CCC GGG AGA	192
	Met Ser Tyr Lys Pro Gly Phe Arg Asp * Ser Gly Leu Pro Gly Arg	
	620 625 630	
30	AGG TTT ATC CTT CGT TGA CCG TGA GAG CTT ATA ATG GGC TAA GTT GGG	240
	Arg Phe Ile Leu Arg * Pro * Glu Leu Ile Met Gly * Val Gly	
	635 640 645	
35	ACA CCC CTG CAG GGT ATT ATC TTT CGA AAG CCG TGC CCG CGG TTA TGA	288
	Thr Pro Leu Gln Gly Ile Ile Phe Arg Lys Pro Cys Pro Arg Leu *	
	650 655 660	
40	GGC AGA TGG GAA TTT GTT AAT GTC CGA TTG TAG AGA ACC TGT CAC TTG	336
	Gly Arg Trp Glu Phe Val Asn Val Arg Leu * Arg Thr Cys His Leu	
	665 670 675 680	
45	ACT TAA TTT AAA ATT CAT CAA CCG TGT GTG TAG CCG TGA TGG TCT CTT	384
	Thr * Phe Lys Ile His Gln Pro Cys Val * Pro * Trp Ser Leu	
	685 690 695	
50	TTC GGC GGA GTC CGG GAA GTG AAC ACG GTT TGA GTT ATG CAT GAA CGT	432
	Phe Gly Gly Val Arg Glu Val Asn Thr Val * Val Met His Glu Arg	
	700 705 710	
55	AAG TAG TTT CAG GAT CAC TCC TTG ATC ACT TCT AGC TCC GCG ACC GTT	480
	Lys * Phe Gln Asp His Ser Leu Ile Thr Ser Ser Ser Ala Thr Val	
	715 720 725	
60	GCG TTG TTT CTC TTC TCG CTC TCA TTT GCG TAT GTT AGC CAC CAT ATA	528
	Ala Leu Phe Leu Phe Ser Leu Ser Phe Ala Tyr Val Ser His His Ile	
	730 735 740	
65	TGC TTA GTG TCT GCT GCA GCT CCA CCT CAT TAC CCC TTC CTT TCC TAT	576
	Cys Leu Val Ser Ala Ala Ala Pro Pro His Tyr Pro Phe Leu Ser Tyr	
	745 750 755 760	
70	AAG CTT AAA TAG TCT TGA TCT CGC GGG TGT GAG ATT GCT GAG TCC TCG	624
	Lys Leu Lys * Ser * Ser Arg Gly Cys Glu Ile Ala Glu Ser Ser	
	765 770 775	

- 104 -

	TGA	CTT	ACA	GAT	TCT	ACC	AAA	ACA	GTT	GCA	GGT	GTC	GAC	GAT	GCC	AGT	672
	*	Leu	Thr	Asp	Ser	Thr	Lys	Thr	Val	Ala	Gly	Val	Asp	Asp	Ala	Ser	
				780					785					790			
5	GCA	GGT	GAC	GCA	ACC	GAG	CTC	AAG	TGG	GAG	TTC	GAC	GAG	GAA	CGT	GGT	720
	Ala	Gly	Asp	Ala	Thr	Glu	Leu	Lys	Trp	Glu	Phe	Asp	Glu	Glu	Arg	Gly	
			795					800					805				
10	CGT	TAC	TAT	GTT	TCT	TTT	CCT	GAT	GAT	CAG	TAG	TGG	AGC	CCA	GTT	GGG	768
	Arg	Tyr	Tyr	Val	Ser	Phe	Pro	Asp	Asp	Gln	*	Trp	Ser	Pro	Val	Gly	
		810					815					820					
15	ACG	ATC	GGG	GAT	CTA	GCA	TTT	GGG	GTT	ATC	TTA	ATT	TCT	TTT	AGA	TTT	816
	Thr	Ile	Gly	Asp	Leu	Ala	Phe	Gly	Val	Ile	Leu	Ile	Ser	Phe	Arg	Phe	
	825					830					835					840	
20	GAC	CGT	AAT	CGG	TCT	ATG	TGT	GGA	TTT	TGG	ATG	ATG	TAT	GAA	TTA	TTT	864
	Asp	Arg	Asn	Arg	Ser	Met	Cys	Gly	Phe	Trp	Met	Met	Tyr	Glu	Leu	Phe	
				845						850					855		
	ATG	TAT	TGT	GTG	AAG	TGG	CGA	TTG	TAA	GCC	AAC	TCT	CGT	TAT	CCC	ATT	912
	Met	Tyr	Cys	Val	Lys	Trp	Arg	Leu	*	Ala	Asn	Ser	Arg	Tyr	Pro	Ile	
				860					865					870			
25	CTT	GTT	CAT	TAC	ATG	GGA	TTG	TGT	GAA	GAT	GAC	CCT	TCT	TGC	GAC	AAA	960
	Leu	Val	His	Tyr	Met	Gly	Leu	Cys	Glu	Asp	Asp	Pro	Ser	Cys	Asp	Lys	
			875					880					885				
30	ACC	ACA	ATG	CGG	TTA	TGC	CTC	TAA	GTC	GTG	CCT	CGA	CAC	GTG	GGA	GAT	1008
	Thr	Thr	Met	Arg	Leu	Cys	Leu	*	Val	Val	Pro	Arg	His	Val	Gly	Asp	
			890				895					900					
35	ATA	GCC	GCA	TCG	TGG	GCG	TTA	CAC	GCA	AGT	CTT	CAT	AGC	AAC	CAA	AAC	1056
	Ile	Ala	Ala	Ser	Trp	Ala	Leu	His	Ala	Ser	Leu	His	Ser	Asn	Gln	Asn	
	905					910					915					920	
40	TCC	TCT	CCG	CAT	TAC	AAG	CCA	CCA	ATC	GCA	GCC	ACC	ATG	ACT	TTC	TTC	1104
	Ser	Ser	Pro	His	Tyr	Lys	Pro	Pro	Ile	Ala	Ala	Thr	Met	Thr	Phe	Phe	
				925						930					935		
	ACC	ACT	GTC	AAT	GCC	ATG	AAA	ATC	TAT	ATG	TAG	ACA	TGT	CCC	ATT	GCA	1152
	Thr	Thr	Val	Asn	Ala	Met	Lys	Ile	Tyr	Met	*	Thr	Cys	Pro	Ile	Ala	
				940					945					950			
45	TCG	GCA	AGA	AAG	CGA	AGC	TTC	ACG	GCA	CAC	CTT	CAT	GAA	GCC	TCT	CTG	1200
	Ser	Ala	Arg	Lys	Arg	Ser	Phe	Thr	Ala	His	Leu	His	Glu	Ala	Ser	Leu	
			955					960					965				
50	GCC	GAA	GAC	AAG	GAT	GCG	CCC	GAC	CGG	ATC	AAT	TCC	TAT	CTA	GAT	ACC	1248
	Ala	Glu	Asp	Lys	Asp	Ala	Pro	Asp	Arg	Ile	Asn	Ser	Tyr	Leu	Asp	Thr	
		970					975					980					
55	TAG	TGG	AGC	CAT	GCG	CCA	ATA	GCG	GAG	ATC	TCC	GAG	AGG	AAG	ACC	GGA	1296
	*	Trp	Ser	His	Ala	Pro	Ile	Ala	Glu	Ile	Ser	Glu	Arg	Lys	Thr	Gly	
	985					990					995					1000	
60	ACT	CGT	CGG	ACG	TCG	GCG	TCC	AAA	TCG	AGG	AGG	CCG	GCA	TGA	AGC	ACA	1344
	Thr	Arg	Arg	Thr	Ser	Ala	Ser	Lys	Ser	Arg	Arg	Pro	Ala	*	Ser	Thr	
				1005						1010					1015		
	TCG	AGG	ATG	GTG	ATC	CCC	ATA	CGG	GTA	GAT	CGG	GTC	GGC	CGC	CAT	CTC	1392
	Ser	Arg	Met	Val	Ile	Pro	Ile	Arg	Val	Asp	Arg	Val	Gly	Arg	His	Leu	
				1020					1025					1030			

- 105 -

	ACA	CCG	AGA	TTA	GGA	TGC	TTA	AAA	CGG	TTT	TTT	TGG	CAC	TAG	CAT	TAT	1440
	Thr	Pro	Arg	Leu	Gly	Cys	Leu	Lys	Arg	Phe	Phe	Trp	His	*	His	Tyr	
			1035					1040					1045				
5	TTT	GCA	TCA	TCC	GTT	GGA	GAG	AAC	ATG	AGA	GAG	CCC	CAT	TTC	TTC	CAC	1488
	Phe	Ala	Ser	Ser	Val	Gly	Glu	Asn	Met	Arg	Glu	Pro	His	Phe	Phe	His	
		1050					1055					1060					
10	GGT	TCT	ACC	TAT	GGG	ATC	TTG	TTC	TGC	TTG	CAA	CCG	GGC	CTC	ACG	GAA	1536
	Gly	Ser	Thr	Tyr	Gly	Ile	Leu	Phe	Cys	Leu	Gln	Pro	Gly	Leu	Thr	Glu	
	1065					1070					1075					1080	
15	AAC	CCG	CGC	CAG	CGG	ACC	CAC	CCC	ATG	CTA	GCA	GGG	CAC	GGC	ACC	CGC	1584
	Asn	Pro	Arg	Gln	Arg	Thr	His	Pro	Met	Leu	Ala	Gly	His	Gly	Thr	Arg	
				1085						1090					1095		
20	AGC	GGC	CGG	TCC	AAA	TGG	ACG	GTG	AGA	ACC	GCA	ACG	CGA	CAC	GCC	CGG	1632
	Ser	Gly	Arg	Ser	Lys	Trp	Thr	Val	Arg	Thr	Ala	Thr	Arg	His	Ala	Arg	
				1100				1105						1110			
	CAC	TGT	CAG	CAA	AGC	GAG	AGC	GCG	CGC	ACG	GCA	CAC	GCA	CGC	TCG	GAC	1680
	His	Cys	Gln	Gln	Ser	Glu	Ser	Ala	Arg	Thr	Ala	His	GCA	Arg	Ser	Asp	
			1115					1120					1125				
25	GAA	CGG	ACG	GTG	CGA	TCG	ATC	CCT	CCC	CCC	TCG	CTC	AAC	CAC	AGT	AGT	1728
	Glu	Arg	Thr	Val	Arg	Ser	Ile	Pro	Pro	Pro	Ser	Leu	Asn	His	Ser	Ser	
		1130					1135					1140					
30	ACC	CTG	CCA	CAC	TAT	CAC	GCA	CGC	ACT	CGA	GTC	ACA	CCT	CCC	ACG	AAG	1776
	Thr	Leu	Pro	His	Tyr	His	Ala	Arg	Thr	Arg	Val	Thr	Pro	Pro	Thr	Lys	
	1145					1150					1155					1160	
35	AAC	CAA	CAG	GAG	GCG	CGG	ATC	CCA	CCG	ATA	AAT	AAC	CCC	GCC	TCG	CCG	1824
	Asn	Gln	Gln	Glu	Ala	Arg	Ile	Pro	Pro	Ile	Asn	Asn	Pro	Ala	Ser	Pro	
				1165						1170					1175		
40	CTC	CTC	CCC	AAA	ATC	AAT	CAC	CGA	TCG	CTC	GGG	GTT	CCC	GGC	ATG	ACG	1872
	Leu	Leu	Pro	Lys	Ile	Asn	His	Arg	Ser	Leu	Gly	Val	Pro	Gly	Met	Thr	
				1180					1185					1190			
	ATG	ATG	GCC	ATG	GCC	AAG	GCG	CCC	TGC	CTC	TGC	GCG	CGC	CCG	TCC	CTC	1920
	Met	Met	Ala	Met	Ala	Lys	Ala	Pro	Cys	Leu	Cys	Ala	Arg	Pro	Ser	Leu	
			1195					1200					1205				
45	GCC	GCG	CGC	GCG	AGG	CGG	CCG	GGG	CCG	GGG	CCG	GCG	CCG	CGC	CTG	CGA	1968
	Ala	Ala	Arg	Ala	Arg	Arg	Pro	Gly	Pro	Gly	Pro	Ala	Pro	Arg	Leu	Arg	
		1210					1215					1220					
50	CGG	TGG	CGA	CCC	AAT	GCG	ACG	GCG	GGG	AAG	GGG	GTC	GGC	GAG	GTG	TGC	2016
	Arg	Trp	Arg	Pro	Asn	Ala	Thr	Ala	Gly	Lys	Gly	Val	Gly	Glu	Val	Cys	
	1225					1230					1235					1240	
55	GCC	GCG	GTT	GTC	GAG	GCG	GCG	ACG	AAG	GCC	GAG	GAT	GAG	GAC	GAC	GAC	2064
	Ala	Ala	Val	Val	Glu	Ala	Ala	Thr	Lys	Ala	Glu	Asp	Glu	Asp	Asp	Asp	
					1245					1250					1255		
60	GAG	GAG	GAG	GCG	GTG	GCG	GAG	GAC	AGG	TAC	GCG	CTC	GGC	GGC	GCG	TGC	2112
	Glu	Glu	Glu	Ala	Val	Ala	Glu	Asp	Arg	Tyr	Ala	Leu	Gly	Gly	Ala	Cys	
				1260				1265					1270				
	AGG	GTG	CTC	GCC	GGA	ATG	CCC	GCG	CCG	CTG	GGC	GCC	ACC	GCG	CTC	GCC	2160
	Arg	Val	Leu	Ala	Gly	Met	Pro	Ala	Pro	Leu	Gly	Ala	Thr	Ala	Leu	Ala	
			1275					1280					1285				

- 106 -

	GGC	GGG	GTC	AAT	TTC	GCC	GTC	TAC	TCC	GGT	GGA	GCC	ACC	GCC	GCG	GCG	2208
	Gly	Gly	Val	Asn	Phe	Ala	Val	Tyr	Ser	Gly	Gly	Ala	Thr	Ala	Ala	Ala	
		1290					1295					1300					
5	CTC	TGC	CTC	TTC	ACG	CCA	GAA	GAT	CTC	AAG	GCG	GTG	GGG	TTG	CCT	CCC	2256
	Leu	Cys	Leu	Phe	Thr	Pro	Glu	Asp	Leu	Lys	Ala	Val	Gly	Leu	Pro	Pro	
	1305					1310					1315					1320	
10	GAG	TAG	AGT	TCA	TCA	GCT	TTG	CGT	GCG	CCG	CGC	GCC	CCC	TTT	TCT	GGC	2304
	Glu	*	Ser	Ser	Ser	Ala	Leu	Arg	Ala	Pro	Arg	Ala	Pro	Phe	Ser	Gly	
					1325					1330					1335		
15	CTG	CGA	TTT	AAG	TTT	TGT	ACT	GGG	GGA	AAT	GCT	GCA	GGA	TAG	GGT	GAC	2352
	Leu	Arg	Phe	Lys	Phe	Cys	Thr	Gly	Gly	Asn	Ala	Ala	Gly	*	Gly	Asp	
				1340					1345					1350			
20	GGA	GGA	GGT	TTC	CCT	TGA	CCC	CCT	GAT	GAA	TCG	GAC	TGG	GAA	CGT	GTG	2400
	Gly	Gly	Gly	Phe	Pro	*	Pro	Pro	Asp	Glu	Ser	Asp	Trp	Glu	Arg	Val	
			1355					1360					1365				
25	GCA	TGT	CTT	CAT	TGA	AGG	CGA	GCT	GCA	CGA	CAT	GCT	TTA	CGG	GTA	CAG	2448
	Ala	Cys	Leu	His	*	Arg	Arg	Ala	Ala	Arg	His	Ala	Leu	Arg	Val	Gln	
	1370						1375					1380					
30	GTT	CGA	CGG	CAC	CTT	TGC	TCC	TCA	CTG	CGG	GCA	CTA	CCT	TGA	TAT	TTC	2496
	Val	Arg	Arg	His	Leu	Cys	Ser	Ser	Leu	Arg	Ala	Leu	Pro	*	Tyr	Phe	
	1385					1390					1395					1400	
35	CAA	TGT	CGT	GGT	GGA	TCC	TTA	TGC	TAA	GGT	GAT	CAT	ACT	TTA	GCT	TTA	2544
	Gln	Cys	Arg	Gly	Gly	Ser	Leu	Cys	*	Gly	Asp	His	Thr	Leu	Ala	Leu	
				1405					1410						1415		
40	CCT	GCA	TCT	TGG	TAT	TTA	CAG	TAG	AAA	TTG	TTA	CGT	GGA	CCC	TTA	TTT	2592
	Pro	Ala	Ser	Trp	Tyr	Leu	Gln	*	Lys	Leu	Leu	Arg	Gly	Pro	Leu	Phe	
			1420					1425						1430			
45	GTT	GCC	TTT	TGT	GTT	GCT	CTA	GGC	AGT	GAT	AAG	CCG	AGG	GGA	GTA	TGG	2640
	Val	Ala	Phe	Cys	Val	Ala	Leu	Gly	Ser	Asp	Lys	Pro	Arg	Gly	Val	Trp	
			1435					1440					1445				
50	CGT	TCC	GGC	GCG	TGG	TAA	CAA	TTG	CTG	GCC	TCA	GAT	GGC	TGG	CAT	GAT	2688
	Arg	Ser	Gly	Ala	Trp	*	Gln	Leu	Leu	Ala	Ser	Asp	Gly	Trp	His	Asp	
		1450				1455						1460					
55	CCC	TCT	TCC	ATA	TAG	CAC	GGT	ATG	CCT	GAT	TGC	TGA	AAA	TAT	TGG	CTG	2736
	Pro	Ser	Ser	Ile	*	His	Gly	Met	Pro	Asp	Cys	*	Lys	Tyr	Trp	Leu	
	1465				1470						1475					1480	
60	CAT	TTG	TTT	CTC	TCT	TTT	TCT	CAT	ATT	TTT	CTC	CTG	TCT	TTC	ACT	TGT	2784
	His	Leu	Phe	Leu	Ser	Phe	Ser	His	Ile	Phe	Leu	Ser	Phe	Thr	Cys		
				1485					1490						1495		
65	ACT	ACA	TTG	CCT	CAG	ACA	GTC	ATG	ATC	AAA	GAG	AGC	AGT	GTC	ATT	AGA	2832
	Thr	Thr	Leu	Pro	Gln	Thr	Val	Met	Ile	Lys	Glu	Ser	Ser	Val	Ile	Arg	
				1500					1505					1510			
70	CAT	TTG	TAG	TTG	TCT	GCT	GAC	TTT	GAC	CAA	AAC	TTG	TAA	TTT	ACT	GTT	2880
	His	Leu	*	Leu	Ser	Ala	Asp	Phe	Asp	Gln	Asn	Leu	*	Phe	Thr	Val	
			1515					1520					1525				
75	GTT	AAA	GGT	CCT	TGA	ATC	ATA	TTT	TTT	TAT	AAT	ATT	ATG	TTT	GCA	AGT	2928
	Val	Lys	Gly	Pro	*	Ile	Ile	Phe	Phe	Tyr	Asn	Ile	Met	Phe	Ala	Ser	
		1530					1535					1540					



- 107 -

	GGA	AGT	AAA	GTG	AAA	TTG	CAT	CTA	GTA	TTT	GTT	GTT	GCT	GTC	TTA	GTC	2976
	Gly	Ser	Lys	Val	Lys	Leu	His	Leu	Val	Phe	Val	Val	Ala	Val	Leu	Val	
	1545					1550					1555					1560	
5	GTT	TAA	TTG	GAC	ATG	CAG	TAA	AAA	GGT	TTG	CAT	CTG	CAG	TTT	GAT	TGG	3024
	Val	*	Leu	Asp	Met	Gln	*	Lys	Gly	Leu	His	Leu	Gln	Phe	Asp	Trp	
					1565					1570					1575		
10	GAA	GGC	GAC	CTA	CCT	CTA	AGA	TAT	CCT	CAA	AAG	GAC	CTG	GTA	ATA	TAT	3072
	Glu	Gly	Asp	Leu	Pro	Leu	Arg	Tyr	Pro	Gln	Lys	Asp	Leu	Val	Ile	Tyr	
				1580					1585						1590		
15	GAG	ATG	CAC	TTG	CGT	GGA	TTC	ACG	AAG	CAT	GAT	TCA	AGC	AAT	GTA	GAA	3120
	Glu	Met	His	Leu	Arg	Gly	Phe	Thr	Lys	His	Asp	Ser	Ser	Asn	Val	Glu	
			1595					1600					1605				
20	CAT	CCG	GGT	ACT	TTC	ATT	GGA	GCT	GTG	TCG	AAG	CTT	GAC	TAT	TTG	AAG	3168
	His	Pro	Gly	Thr	Phe	Ile	Gly	Ala	Val	Ser	Lys	Leu	Asp	Tyr	Leu	Lys	
		1610					1615						1620				
25	GTA	CAG	CTG	TAC	TTG	CTG	ACT	ACA	TAG	GAT	AAT	TTT	TAA	AGA	AAG	CTA	3216
	Val	Gln	Leu	Tyr	Leu	Leu	Thr	Thr	*	Asp	Asn	Phe	*	Arg	Lys	Leu	
	1625					1630					1635					1640	
30	CAT	ATT	AGC	CAG	AAT	TTG	GGT	TAT	TAC	AAA	AAC	TAC	TGC	ATA	CTA	TAG	3264
	His	Ile	Ser	Gln	Asn	Leu	Gly	Tyr	Tyr	Lys	Asn	Tyr	Cys	Ile	Leu	*	
					1645					1650					1655		
35	CAG	TTA	CAT	GCT	CAT	TAT	CGA	GGA	GAT	GCT	CAC	ACG	CAT	CTT	ATT	TGG	3312
	Gln	Leu	His	Ala	His	Tyr	Arg	Gly	Asp	Ala	His	Thr	His	Leu	Ile	Trp	
				1660					1665					1670			
40	ATT	TAA	TAC	CCA	ATT	CTG	TTT	TGA	TAT	TGG	ACT	GTT	CCC	TCT	ACA	GGA	3360
	Ile	*	Tyr	Pro	Ile	Leu	Phe	*	Tyr	Trp	Thr	Val	Pro	Ser	Thr	Gly	
			1675				1680						1685				
45	GCT	TGG	AGT	TAA	TTG	TAT	TGA	ATT	AAT	GCC	CTG	CCA	TGA	GTT	CAA	CGA	3408
	Ala	Trp	Ser	*	Leu	Tyr	*	Ile	Asn	Ala	Leu	Pro	*	Val	Gln	Arg	
	1690						1695					1700					
50	GCT	GGA	GTA	CTC	AAC	CTC	TTC	TTC	CAA	GTA	AGG	ACA	TGA	ATT	TAG	TAT	3456
	Ala	Gly	Val	Leu	Asn	Leu	Phe	Phe	Gln	Val	Arg	Thr	*	Ile	*	Tyr	
	1705					1710					1715					1720	
55	TAG	CCT	GCC	AGC	ACT	GTT	TGA	GTG	AGA	GTT	CAT	ACA	CAT	TTT	GTG	CCT	3504
	*	Pro	Ala	Ser	Thr	Val	*	Val	Arg	Val	His	Thr	His	Phe	Val	Pro	
					1725					1730					1735		
60	GCA	TAA	CTG	ATA	TTT	GTT	CAA	ACT	ATT	TTT	TTT	AGC	AGT	CAC	TCA	ACA	3552
	Ala	*	Leu	Ile	Phe	Val	Gln	Thr	Ile	Phe	Phe	Ser	Ser	His	Ser	Thr	
				1740					1745						1750		
65	GTT	TTA	CAT	ATA	TAT	ATA	ATA	TAG	ACT	ATT	CGT	CAC	CCT	GGG	TGA	GGA	3600
	Val	Leu	His	Ile	Tyr	Ile	Ile	*	Thr	Ile	Arg	His	Pro	Gly	*	Gly	
			1755					1760					1765				
70	ATA	GTT	ATT	CTT	CAC	CCA	CCT	CTA	TTT	TAA	CAT	CTA	TGC	ACC	GTA	ATT	3648
	Ile	Val	Ile	Leu	His	Pro	Pro	Leu	Phe	*	His	Leu	Cys	Thr	Val	Ile	
		1770					1775					1780					
75	TTA	CGT	TTC	GTA	AAT	TTG	TCT	TAT	TTT	AGA	GAT	AAA	AAG	AGA	ACG	TAA	3696
	Leu	Arg	Phe	Val	Asn	Leu	Ser	Tyr	Phe	Arg	Asp	Lys	Lys	Arg	Thr	*	
	1785					1790					1795					1800	

- 108 -

	GAA Glu	AAC Asn	CTA Leu	TAA *	TCG Ser 1805	TCG Ser	TAA *	AAA Lys	AAA Lys	ATA Ile 1810	TGT Cys	TAC Tyr	GTA Val	AAA Lys	TTA Leu 1815	CAA Gln	3744
5	ATG Met	TAA *	AAA Lys	CAT His 1820	AGT Ser	GTA Val	AAA Lys	TGT Cys	ACA Thr 1825	TAA *	AAT Asn	ACA Thr	TTT Phe 1830	TTT Phe	GAC Asp	CTA Leu	3792
10	TAT Tyr	TTT Phe	TTT Phe	TGT Cys 1835	TAA *	TGC Cys	CAA Gln	ATT Ile 1840	TTA Leu	TAC Tyr	AGT Ser	AAA Lys	TCA Ser 1845	ATA Ile	TGA *	ATG Met	3840
15	TAA *	CTA Leu 1850	TTT Phe	GTA Val	TTT Phe	CAA Gln	ATG Met 1855	TAA *	TTT Phe	ATT Ile	TAT Tyr	GAA Glu 1860	ATG Met	GTC Val	GTA Val	AGA Arg	3888
20	TTA Leu 1865	CCT Pro	CGG Arg	GTG Val	AAG Lys	AAT Asn 1870	AAC Asn	TTA Leu	TTC Phe	TGC Cys 1875	ACC Thr	CTG Leu	GGT Gly	GAT Asp	GAA Glu 1880	TAG *	3936
	TAA *	CAC His	TAT Tyr	ATA Ile 1885	TAT Tyr	ATA Ile	TAT Tyr	ATA Ile 1890	TAT Tyr	ATA Ile	TAT Tyr	ATA Ile	TAT Tyr	ATA Ile	CCG Pro 1895	GCT Ala	3984
25	GCT Ala	GCT Ala	AAT Asn	GAT Asp 1900	GTT Val	AAT Asn	ATT Ile	TCG Ser	CAA Gln 1905	GTA Val	CCT Pro	AAG Lys	CTG Leu 1910	GAT Asp	TTT Phe	TCT Ser	4032
30	CCA Pro	TGA *	GAC Asp 1915	ATC Ile	AAT Asn	CCA Pro	TAA *	TTG Leu 1920	AAA Lys	TTG Leu	GTC Val	ACG Thr	ACA Thr 1925	GTT Val	GAA Glu	TAG *	4080
35	TTG Leu 1930	ATA Ile	GCT Ala	GAA Glu	AAT Asn	GAA Glu	ATC Ile 1935	CAG Gln	CAT His	GCT Ala	ACT Thr	GTC Val 1940	TTG Leu	CCA Pro	TCT Ser	CCA Pro	4128
40	GAC Asp 1945	TTG Leu	CTA Leu	ACA Thr	TGA *	ATT Ile 1950	TTG Leu	TCT Ser	GCC Ala	TAC Tyr	CTG Leu 1955	TCA Ser	TTT Phe	GTA Val	CCA Pro	ACG Thr 1960	4176
	TTC Phe	CCA Pro	ATT Ile	GCC Ala	CTC Leu 1965	TCA Ser	TTA Leu	TTC Phe	GTG Val 1970	TGT Cys	ACC Thr	ATG Met	CAT His	ATG Met	TGT Cys 1975	TTT Phe	4224
45	AAC Asn	ATG Met	ATT Ile	ATT Ile 1980	GTT Val	GGC Gly	TAT Tyr	ATT Ile	TCT Ser 1985	CTT Leu	TGG Trp	AAA Lys	CAT His 1990	GAC Asp	TAA *	TTT Phe	4272
50	ATC Ile	ACC Thr	CGT Arg 1995	TTT Phe	GTA Val	TAA *	ACT Thr	GCT Ala 2000	TGT Cys	TTT Phe	CAT His	ATC Ile 2005	AGG Arg 2005	ATG Met	AAC Asn	TTT Phe	4320
55	TGG Trp 2010	GGA Gly	TAT Tyr	TCT Ser	ACC Thr	ATA Ile	AAC Asn 2015	TTC Phe	TTT Phe	TCA Ser	CCA Pro	ATG Met 2020	ACG Thr	AGA Arg	TAC Tyr	ACA Thr	4368
	TCA Ser 2025	GGC Gly	GGG Gly	ATA Ile	AAA Lys	AAC Asn 2030	TGT Cys	GGG Gly	CGT Arg	GAT Asp 2035	GCC Ala	ATA Ile	AAT Asn	GAG Glu	TTC Phe	AAA Lys 2040	4416
60	ACT Thr	TTT Phe	GTA Val	AGA Arg 2045	GAG Glu	GCT Ala	CAC His	AAA Lys	CGG Arg 2050	GGA Gly	ATT Ile 2050	GAG Glu	GTA Val	AGC Ser	AAG Lys 2055	TCG Ser	4464

- 109 -

	TAC	GAG	TTA	GTT	GCT	CCT	TTT	GAA	CTT	ATC	AAT	TTG	ATG	CGA	AGA	CAT	4512
	Tyr	Glu	Leu	Val	Ala	Pro	Phe	Glu	Leu	Ile	Asn	Leu	Met	Arg	Arg	His	
				2060					2065					2070			
5	GTT	ACT	GCT	AGG	TGA	TCC	TGG	ATG	TTG	TCT	TCA	ACC	ATA	CAG	CTG	AGG	4560
	Val	Thr	Ala	Arg	*	Ser	Trp	Met	Leu	Ser	Ser	Thr	Ile	Gln	Leu	Arg	
			2075					2080					2085				
10	GTA	ATG	AGA	ATG	GTC	CAA	TAT	TAT	CAT	TTA	GGG	GGG	TCG	ATA	ATA	CTA	4608
	Val	Met	Arg	Met	Val	Gln	Tyr	Tyr	His	Leu	Gly	Gly	Ser	Ile	Ile	Leu	
		2090					2095					2100					
15	CAT	ACT	ATA	TGC	TTG	CAC	CCA	AGG	TGA	CAG	ATC	TTT	CTT	GCT	GCG	TAA	4656
	His	Thr	Ile	Cys	Leu	His	Pro	Arg	*	Gln	Ile	Phe	Leu	Ala	Ala	*	
	2105					2110					2115					2120	
20	TTG	TTC	TTT	CAT	AGA	TGT	ATA	GAG	CAT	AGA	TGT	GTT	ATG	TAG	TAG	TTC	4704
	Leu	Phe	Phe	His	Arg	Cys	Ile	Glu	His	Arg	Cys	Val	Met	*	*	Phe	
				2125						2130					2135		
	TTT	TTC	AAG	GGG	ATT	ATG	TTC	ATG	CAG	GGA	GAG	TTT	TAT	AAC	TAT	TCT	4752
	Phe	Phe	Lys	Gly	Ile	Met	Phe	Met	Gln	Gly	Glu	Phe	Tyr	Asn	Tyr	Ser	
				2140					2145					2150			
25	GGC	TGT	GGG	AAT	ACC	TTC	AAC	TGT	AAT	CAT	CCT	GTG	GTT	CGT	CAA	TTC	4800
	Gly	Cys	Gly	Asn	Thr	Phe	Asn	Cys	Asn	His	Pro	Val	Val	Arg	Gln	Phe	
			2155					2160					2165				
30	ATT	GTA	GAT	TGT	TTA	AGG	TAC	AGA	TAT	ACA	TTT	TAC	TTC	TAG	AAC	TAC	4848
	Ile	Val	Asp	Cys	Leu	Arg	Tyr	Arg	Tyr	Thr	Phe	Tyr	Phe	*	Asn	Tyr	
		2170					2175					2180					
35	TTT	TTC	ATT	TCT	TTT	GCT	GCT	TGT	CAT	TTT	GAT	ATG	ATT	AAT	TTG	CAA	4896
	Phe	Phe	Ile	Ser	Phe	Ala	Ala	Cys	His	Phe	Asp	Met	Ile	Asn	Leu	Gln	
	2185					2190					2195					2200	
40	GCT	TGT	GGG	GGT	AAA	TCT	TTT	GGT	CAG	CAT	ATT	GTA	TCT	TTA	AAT	GTC	4944
	Ala	Cys	Gly	Gly	Lys	Ser	Phe	Gly	Gln	His	Ile	Val	Ser	Leu	Asn	Val	
				2205						2210					2215		
	ACA	AAT	ACT	AAT	GTC	CTG	GTG	CTT	ATT	GAT	TTG	GCA	TCT	TCA	AAT	TCT	4992
	Thr	Asn	Thr	Asn	Val	Leu	Val	Leu	Ile	Asp	Leu	Ala	Ser	Ser	Asn	Ser	
				2220					2225					2230			
45	TCT	CCA	ATG	AAA	AGG	GAA	AAA	TCT	ACT	GTA	TGT	CTC	GTC	AAC	TAA	TTT	5040
	Ser	Pro	Met	Lys	Arg	Glu	Lys	Ser	Thr	Val	Cys	Leu	Val	Asn	*	Phe	
			2235					2240					2245				
50	ACT	TTT	GTT	TTG	CAG	ATA	CTG	GGT	GAT	GGA	AAT	GCA	TGT	TGA	TGG	TTT	5088
	Thr	Phe	Val	Leu	Gln	Ile	Leu	Gly	Asp	Gly	Asn	Ala	Cys	*	Trp	Phe	
		2250					2255					2260					
55	TCG	TTT	TGA	TCT	TGC	ATC	CAT	AAT	GAC	CAG	AGG	TTC	CAG	GTA	ATT	TGT	5136
	Ser	Phe	*	Ser	Cys	Ile	His	Asn	Asp	Gln	Arg	Phe	Gln	Val	Ile	Cys	
	2265					2270					2275					2280	
60	ATT	TAT	TGT	TTG	TTT	GCG	TGT	TGC	CTT	TTC	AGA	AGA	TTC	TTA	AAA	GAA	5184
	Ile	Tyr	Cys	Leu	Phe	Ala	Cys	Cys	Leu	Phe	Arg	Arg	Phe	Leu	Lys	Glu	
				2285					2290						2295		
	TGT	TTC	TTT	TAC	AAG	TCT	GTG	GGA	TCC	AGT	TAA	CGT	GTA	TGG	AGC	TCC	5232
	Cys	Phe	Phe	Tyr	Lys	Ser	Val	Gly	Ser	Ser	*	Arg	Val	Trp	Ser	Ser	
				2300					2305					2310			

- 110 -

	AAT	AGA	AGG	TGA	CAT	GAT	CAC	AAC	AGG	GAC	ACC	TCT	TGT	TAC	TCC	ACC	5280
	Asn	Arg	Arg	*	His	Asp	His	Asn	Arg	Asp	Thr	Ser	Cys	Tyr	Ser	Thr	
			2315						2320				2325				
5	ACT	TAT	TGA	CAT	GAT	CAG	CAA	TGA	CCC	AAT	TCT	TGG	AGG	CGT	CAA	GGT	5328
	Thr	Tyr	*	His	Asp	Gln	Gln	*	Pro	Asn	Ser	Trp	Arg	Arg	Gln	Gly	
		2330					2335					2340					
10	ACT	TGT	TTC	ATC	CAA	CAC	CTG	TTG	TCT	GTG	TGC	ATT	CAA	TTG	TTT	TAA	5376
	Thr	Cys	Phe	Ile	Gln	His	Leu	Leu	Ser	Val	Cys	Ile	Gln	Leu	Phe	*	
		2345				2350					2355					2360	
15	TAT	GGT	AAT	GAT	CAA	TTT	CCC	AAT	GTT	GAT	AAG	GAA	AAA	AAA	TGC	AAG	5424
	Tyr	Gly	Asn	Asp	Gln	Phe	Pro	Asn	Val	Asp	Lys	Glu	Lys	Lys	Cys	Lys	
				2365						2370					2375		
20	TAG	CTC	TCT	TTA	TCT	GCT	TCT	TGT	GAG	TTA	TGC	TAA	ACA	TGT	AGA	TAC	5472
	*	Leu	Ser	Leu	Ser	Ala	Ser	Cys	Glu	Leu	Cys	*	Thr	Cys	Arg	Tyr	
				2380					2385					2390			
25	TAC	TAT	ATT	TCA	ACT	GTA	TAT	ACT	TGA	CAT	ATT	ATT	GCT	TCC	TTG	GGA	5520
	Tyr	Tyr	Ile	Ser	Thr	Val	Tyr	Thr	*	His	Ile	Ile	Ala	Ser	Leu	Gly	
			2395					2400					2405				
30	GGC	TCT	CTT	ATT	CCT	TTC	CCC	CGT	TGC	AAT	TAT	AGC	TCA	TTG	CTG	AAG	5568
	Gly	Ser	Leu	Ile	Pro	Phe	Pro	Arg	Cys	Asn	Tyr	Ser	Ser	Leu	Leu	Lys	
		2410					2415					2420					
35	CAT	GGG	ATG	CAG	GAG	GCC	TCT	ATC	AAG	TAG	GTC	AAT	TCC	CTC	ACT	GGA	5616
	His	Gly	Met	Gln	Glu	Ala	Ser	Ile	Lys	*	Val	Asn	Ser	Leu	Thr	Gly	
		2425				2430					2435					2440	
40	ATG	TTT	GGT	CTG	AGT	GGA	ATG	GGA	AGG	TAA	GGT	ACC	TGT	TAA	AAG	TTT	5664
	Met	Phe	Gly	Leu	Ser	Gly	Met	Gly	Arg	*	Gly	Thr	Cys	*	Lys	Phe	
				2445						2450					2455		
45	GAA	TGG	CAA	ATA	CTG	ATA	GAA	ATA	TAA	CTT	ATA	TTT	GCG	ACA	TAT	ATA	5712
	Glu	Trp	Gln	Ile	Leu	Ile	Glu	Ile	*	Leu	Ile	Phe	Ala	Thr	Tyr	Ile	
				2460					2465					2470			
50	GAT	AAA	GCA	AAA	TAA	TAC	GCA	TTC	CAC	CTG	AAC	TTT	AAA	GGG	GCA	CGC	5760
	Asp	Lys	Ala	Lys	*	Tyr	Ala	Phe	His	Leu	Asn	Phe	Lys	Gly	Ala	Arg	
			2475					2480					2485				
55	AGA	ATT	ATC	CCG	CAT	CTG	TCT	ACA	AGA	ATG	ATA	ACA	CAT	GTG	CTG	AAT	5808
	Arg	Ile	Ile	Pro	His	Leu	Ser	Thr	Arg	Met	Ile	Thr	His	Val	Leu	Asn	
		2490					2495					2500					
60	AGT	GAA	GTA	CTA	CTT	CTC	AAA	TGT	CTG	AAT	GAA	CGC	ACT	AAC	TCT	TGT	5856
	Ser	Glu	Val	Leu	Leu	Leu	Lys	Cys	Leu	Asn	Glu	Arg	Thr	Asn	Ser	Cys	
		2505				2510					2515					2520	
65	GAG	TGT	CAA	CCG	AGC	AAG	AAA	TAT	TTG	AGT	TTT	CTG	CAA	GAA	ATT	GTT	5904
	Glu	Cys	Gln	Pro	Ser	Lys	Lys	Tyr	Leu	Ser	Phe	Leu	Gln	Glu	Ile	Val	
				2525						2530					2535		
70	CAT	GTT	GTG	CTG	TAT	TAT	ACT	CCC	TCC	GTC	CGA	AAT	TAT	TTG	TCG	GAG	5952
	His	Val	Val	Leu	Tyr	Tyr	Thr	Pro	Ser	Val	Arg	Asn	Tyr	Leu	Ser	Glu	
				2540					2545					2550			
75	AAA	TGG	ATG	TAT	CTA	GAC	GTA	TTT	TAG	TTC	TAG	ATA	CAT	CCA	TTT	TTA	6000
	Lys	Trp	Met	Tyr	Leu	Asp	Val	Phe	*	Phe	*	Ile	His	Pro	Phe	Leu	
			2555					2560					2565				

- 111 -

	TCC	ATT	TCT	GCA	ACA	AGT	AGT	TCC	GGA	CGG	AGG	GAG	TAT	CAT	TTA	ACA	6048
	Ser	Ile	Ser	Ala	Thr	Ser	Ser	Ser	Gly	Arg	Arg	Glu	Tyr	His	Leu	Thr	
	2570							2575				2580					
5	AAT	ATA	TGC	ATG	TTC	GAA	GTA	AAT	CCC	CAC	GAA	TAA	GCA	TAT	AAG	ACG	6096
	Asn	Ile	Cys	Met	Phe	Glu	Val	Asn	Pro	His	Glu	*	Ala	Tyr	Lys	Thr	
	2585					2590					2595					2600	
10	ATA	TTG	CTT	TTT	GAC	TTG	CAA	CAC	CTA	AAC	CTC	ATT	GTT	TTC	TCC	TAG	6144
	Ile	Leu	Leu	Phe	Asp	Leu	Gln	His	Leu	Asn	Leu	Ile	Val	Phe	Ser	*	
					2605					2610					2615		
15	GAT	TTT	GGG	TGT	TCG	AAG	CAA	GCA	GCT	GGT	GAT	ATT	TAA	TTT	ACC	TTT	6192
	Asp	Phe	Gly	Cys	Ser	Lys	Gln	Ala	Ala	Gly	Asp	Ile	*	Phe	Thr	Phe	
				2620					2625					2630			
20	GCC	TTT	ATT	TGT	AGC	TTG	ATT	TGA	GGG	TGC	GGC	AAA	GGT	TTT	AGC	TTA	6240
	Ala	Phe	Ile	Cys	Ser	Leu	Ile	*	Gly	Cys	Gly	Lys	Gly	Phe	Ser	Leu	
			2635					2640					2645				
25	GTA	GTG	TTT	TGT	AAA	TTA	TTA	TAG	TTT	ATG	TAT	ATA	CTC	CTC	ATT	TGG	6288
	Val	Val	Phe	Cys	Lys	Leu	Leu	*	Phe	Met	Tyr	Ile	Leu	Leu	Ile	Trp	
		2650					2655					2660					
30	GCA	CTT	CCG	TAC	TGG	TCC	CAT	AGA	AGA	TAA	AAA	TGG	AAT	GAT	GTC	TGG	6336
	Ala	Leu	Pro	Tyr	Trp	Ser	His	Arg	Arg	*	Lys	Trp	Asn	Asp	Val	Trp	
	2665					2670					2675					2680	
35	CCA	ATA	ATT	GTT	GAC	AAC	ACT	GTT	GCG	CAT	TTG	ATT	TTT	ATC	AGG	GAA	6384
	Pro	Ile	Ile	Val	Asp	Asn	Thr	Val	Ala	His	Leu	Ile	Phe	Ile	Arg	Glu	
					2685					2690					2695		
40	TGG	AAA	ATT	GAA	ATC	GGT	AAG	AAA	CAT	TGC	GAT	ATT	AAG	CTT	GTA	TAT	6432
	Trp	Lys	Ile	Glu	Ile	Gly	Lys	Lys	His	Cys	Asp	Ile	Lys	Leu	Val	Tyr	
			2700						2705					2710			
45	GCT	AAT	GCT	GGT	GGA	TCT	TTA	AGA	GGG	AAC	ATA	TGA	TCT	CGT	GTG	CAT	6480
	Ala	Asn	Ala	Gly	Gly	Ser	Leu	Arg	Gly	Asn	Ile	*	Ser	Arg	Val	His	
			2715					2720					2725				
50	CCA	TCT	TCA	ACT	AAA	AAA	ATA	TGT	TGC	ACA	TCT	CCC	ACG	TCA	CTT	ACT	6528
	Pro	Ser	Ser	Thr	Lys	Lys	Ile	Cys	Cys	Thr	Ser	Pro	Thr	Ser	Leu	Thr	
		2730					2735					2740					
55	AGC	TAT	TTC	ATC	CAA	GTA	CTA	ACT	TGT	GTG	GTT	GTC	TCC	TCA	GTA	CCG	6576
	Ser	Tyr	Phe	Ile	Gln	Val	Leu	Thr	Cys	Val	Val	Val	Ser	Ser	Val	Pro	
	2745					2750					2755					2760	
60	GGA	CAT	TGT	GCG	CCA	ATT	CAT	TAA	AGG	CAC	TGA	TGG	ATT	TGC	TGG	TGG	6624
	Gly	His	Cys	Ala	Pro	Ile	His	*	Arg	His	*	Trp	Ile	Cys	Trp	Trp	
				2765					2770						2775		
65	TTT	TGC	CGA	ATG	TCT	TTG	TGG	AAG	TCC	ACA	CCT	ATA	CCA	GGT	AAG	TTG	6672
	Phe	Cys	Arg	Met	Ser	Leu	Trp	Lys	Ser	Thr	Pro	Ile	Pro	Gly	Lys	Leu	
				2780					2785					2790			
70	TGG	CAA	TAC	TTG	GAA	ATG	GGT	TGA	GTG	AAT	GTC	ACA	TGG	ATT	TTT	TAT	6720
	Trp	Gln	Tyr	Leu	Glu	Met	Gly	*	Val	Asn	Val	Thr	Trp	Ile	Phe	Tyr	
			2795				2800						2805				
75	ATA	TAC	CAC	ATG	ATG	ATA	CAC	ATG	TAA	ATA	TAT	AAC	GAT	TAT	AGT	GTA	6768
	Ile	Tyr	His	Met	Met	Ile	His	Met	*	Ile	Tyr	Asn	Asp	Tyr	Ser	Val	
		2810					2815						2820				

- 112 -

	TGC	ATA	TGC	ATT	TGG	CTA	AGA	AGT	ACT	CCC	TCC	CTT	AGT	AAA	AGT	TAG	6816
	Cys	Ile	Cys	Ile	Trp	Leu	Arg	Ser	Thr	Pro	Ser	Leu	Ser	Lys	Ser	*	
	2825					2830					2835					2840	
5	TAC	AAA	GTT	GAG	TCA	TCT	ATT	TTG	GAA	CGG	AGG	GAG	TAT	AAG	TGT	ATA	6864
	Tyr	Lys	Val	Glu	Ser	Ser	Ile	Leu	Glu	Arg	Arg	Glu	Tyr	Lys	Cys	Ile	
					2845					2850					2855		
10	CAC	TAG	TGC	AAT	ATA	TAG	GTT	TTA	ACA	CCC	AAC	TTG	CCA	ATG	AAG	GAA	6912
	His	*	Cys	Asn	Ile	*	Val	Leu	Thr	Pro	Asn	Leu	Pro	Met	Lys	Glu	
				2860					2865					2870			
15	CAT	AGG	GCT	TTC	TAG	TTA	TCT	TAT	TTA	TTT	GTC	TGG	TGA	ATA	ATC	CAC	6960
	His	Arg	Ala	Phe	*	Leu	Ser	Tyr	Leu	Phe	Val	Trp	*	Ile	Ile	His	
			2875					2880					2885				
20	TGA	AAA	ATT	CCA	GCC	ATG	TCA	TTT	TTT	AGG	GGG	GGA	GAA	GAA	ACT	ACA	7008
	*	Lys	Ile	Pro	Ala	Met	Ser	Phe	Phe	Arg	Gly	Gly	Glu	Glu	Thr	Thr	
		2890					2895					2900					
25	TTG	ATT	TTT	CCC	CCT	AAA	AAA	AGC	CAT	CTC	AGA	TTT	CAT	AGG	TAA	CTT	7056
	Leu	Ile	Phe	Pro	Pro	Lys	Lys	Ser	His	Leu	Arg	Phe	His	Arg	*	Leu	
	2905					2910					2915					2920	
30	GCT	TTT	CTG	TAA	AGA	AAT	GAA	AAC	GAC	TTC	ATA	CTT	TCT	GTC	GAT	TAT	7104
	Ala	Phe	Leu	*	Arg	Asn	Glu	Asn	Asp	Phe	Ile	Leu	Ser	Val	Asp	Tyr	
					2925					2930					2935		
35	AAG	TGT	ATA	CAC	TAG	TGC	AAT	ATA	TAG	GTT	TTA	ACA	CCC	AAC	TTG	CCA	7152
	Lys	Cys	Ile	His	*	Cys	Asn	Ile	*	Val	Leu	Thr	Pro	Asn	Leu	Pro	
				2940					2945					2950			
40	ATG	AAG	GAA	CAT	AGG	GCT	TTC	TAG	TTA	TCT	TAT	TTA	TTT	GCT	GGT	GAA	7200
	Met	Lys	Glu	His	Arg	Ala	Phe	*	Leu	Ser	Tyr	Leu	Phe	Ala	Gly	Glu	
			2955				2960						2965				
45	TAA	TCC	ACT	GAA	AAA	TTC	CAG	CCA	TGT	CAT	TTT	TTA	GGG	GGG	AGA	AGA	7248
	*	Ser	Thr	Glu	Lys	Phe	Gln	Pro	Cys	His	Phe	Leu	Gly	Gly	Arg	Arg	
		2970					2975					2980					
50	AAC	TAT	ATT	GAT	TTT	TCC	CCC	TAA	AAA	AAG	CCA	TCT	CAG	ATT	CAT	AGG	7296
	Asn	Tyr	Ile	Asp	Phe	Ser	Pro	*	Lys	Lys	Pro	Ser	Gln	Ile	His	Arg	
	2985					2990					2995					3000	
55	AAC	TTG	CTT	TTC	TGT	AAA	GAA	ATG	AAA	ACG	ACT	TCA	TAC	TTT	CTG	CGG	7344
	Asn	Leu	Leu	Phe	Cys	Lys	Glu	Met	Lys	Thr	Thr	Ser	Tyr	Phe	Leu	Arg	
				3005						3010					3015		
60	CGC	TTA	CTT	AGC	TCG	ATG	GAT	ATT	TGT	AAG	ATG	AAT	GCC	AAA	TTA	TTT	7392
	Arg	Leu	Leu	Ser	Ser	Met	Asp	Ile	Cys	Lys	Met	Asn	Ala	Lys	Leu	Phe	
				3020					3025					3030			
65	GGC	GGG	ATT	TGA	TCG	TTA	TTC	CAA	ATT	TCA	TTT	GGT	TTC	TCT	AGC	AAT	7440
	Gly	Gly	Ile	*	Ser	Leu	Phe	Gln	Ile	Ser	Phe	Gly	Phe	Ser	Ser	Asn	
			3035					3040					3045				
70	CAA	CCC	AGT	ACC	TTG	TTA	TTG	GCA	CTG	CAA	TTT	CTT	ATT	GAT	TAA	TCA	7488
	Gln	Pro	Ser	Thr	Leu	Leu	Leu	Ala	Leu	Gln	Phe	Leu	Ile	Asp	*	Ser	
		3050					3055					3060					
75	GGC	AGG	AGG	AAG	GAA	ACC	TTG	GCA	CAG	TAT	CAA	CTT	GGT	ATG	TGC	ACA	7536
	Gly	Arg	Arg	Lys	Glu	Thr	Leu	Ala	Gln	Tyr	Gln	Leu	Gly	Met	Cys	Thr	
	3065					3070					3075					3080	

- 113 -

	TGA	TGG	ATT	TAC	ACT	GGG	TGA	TTT	GGT	ACA	TAT	AAT	ACC	AAG	TCA	ATT	7584
	*	Trp	Ile	Tyr	Thr	Gly	*	Phe	Gly	Thr	Tyr	Asn	Thr	Lys	Ser	Ile	
					3085					3090					3095		
5	TAC	CAA	ATG	GGG	AGA	CCA	ATA	GAG	ATG	GAG	AAA	ATC	ACA	ATC	TTA	GCT	7632
	Tyr	Gln	Met	Gly	Arg	Pro	Ile	Glu	Met	Glu	Lys	Ile	Thr	Ile	Leu	Ala	
				3100				3105						3110			
10	GGA	ATT	GTG	GGG	AGG	TAA	TTC	TGA	ACT	CTC	CTT	TTT	TTT	TGA	AAT	TTT	7680
	Gly	Ile	Val	Gly	Arg	*	Phe	*	Thr	Leu	Leu	Phe	*	Asn	Phe		
			3115					3120						3125			
15	CAT	GCT	TTA	CAT	AAT	AGT	CAA	ATG	GCT	GAC	AAA	TGT	CGT	TGT	ATG	GTT	7728
	His	Ala	Leu	His	Asn	Ser	Gln	Met	Ala	Asp	Lys	Cys	Arg	Cys	Met	Val	
		3130					3135						3140				
20	CTC	TCT	ACC	TAA	ACC	GTT	AAG	GCA	GTA	AGA	GTT	TCC	CTA	CAA	GAT	CTC	7776
	Leu	Ser	Thr	*	Thr	Val	Lys	Ala	Val	Arg	Val	Ser	Leu	Gln	Asp	Leu	
	3145					3150					3155					3160	
	TTT	GTT	CGT	ATA	ATT	GTA	TTT	TCT	AGA	GAA	AAG	TTG	CCT	TCA	ATT	TTG	7824
	Phe	Val	Arg	Ile	Ile	Val	Phe	Ser	Arg	Glu	Lys	Leu	Pro	Ser	Ile	Leu	
				3165						3170					3175		
25	TGC	ACG	CGG	CAG	TAC	AGG	AAT	TGT	GGT	TAT	AAA	TAT	TGA	TAC	AGG	CTG	7872
	Cys	Thr	Arg	Gln	Tyr	Arg	Asn	Cys	Gly	Tyr	Lys	Tyr	*	Tyr	Arg	Leu	
				3180				3185						3190			
30	ACC	ATC	GTT	ACT	AAT	AGG	GGG	AAC	AAT	AAG	CAC	ATT	TTT	TTA	ATA	GCA	7920
	Thr	Ile	Val	Thr	Asn	Arg	Gly	Asn	Asn	Lys	His	Ile	Phe	Leu	Ile	Ala	
			3195				3200						3205				
35	AAG	GCA	TCA	CCC	TTG	TTC	CGT	TTC	CAA	TGA	AAT	CAC	AGT	ATC	CGA	ACC	7968
	Lys	Ala	Ser	Pro	Leu	Phe	Arg	Phe	Gln	*	Asn	His	Ser	Ile	Arg	Thr	
		3210					3215						3220				
40	ATA	AGT	TTT	ACA	AGT	ATG	CGT	AGA	GAG	AAA	TAA	AGT	ATC	AAC	CCG	GCA	8016
	Ile	Ser	Phe	Thr	Ser	Met	Arg	Arg	Glu	Lys	*	Ser	Ile	Asn	Pro	Ala	
	3225					3230					3235					3240	
	GAA	ACA	GTT	GTT	TCA	GGC	GCA	AAG	AGA	AAA	GGA	AAC	GAT	ATG	CTC	TAT	8064
	Glu	Thr	Val	Val	Ser	Gly	Ala	Lys	Arg	Lys	Gly	Asn	Asp	Met	Leu	Tyr	
				3245				3250							3255		
45	TAC	ATC	AAC	CTT	TTA	GCA	TTT	AGG	GAC	GAC	CAG	CAT	CAT	CCC	ATC	TTC	8112
	Tyr	Ile	Asn	Leu	Leu	Ala	Phe	Arg	Asp	Asp	Gln	His	His	Pro	Ile	Phe	
				3260				3265						3270			
50	AAT	CAA	CTG	GAG	CGA	GGT	CAC	CTC	CAA	TCT	TCT	CAG	CAG	CCT	CAG	AGT	8160
	Asn	Gln	Leu	Glu	Arg	Gly	His	Leu	Gln	Ser	Ser	Gln	Gln	Pro	Gln	Ser	
			3275					3280						3285			
55	GGT	GAC	CTC	CCA	AGC	AAG	TGC	ATC	AGC	ATC	CAT	CAT	CTG	GGG	GTT	GGG	8208
	Gly	Asp	Leu	Pro	Ser	Lys	Cys	Ile	Ser	Ile	His	His	Leu	Gly	Val	Gly	
		3290					3295						3300				
60	CAC	ATA	CCA	TGA	GCA	CAA	TCA	CCT	GAA	TTT	GAT	GAA	TTT	TCC	TCT	GTT	8256
	His	Ile	Pro	*	Ala	Gln	Ser	Pro	Glu	Phe	Asp	Glu	Phe	Ser	Ser	Val	
	3305					3310					3315					3320	
	TAC	CTT	GCA	GCA	GAC	CCC	TGC	CGT	ATA	AAT	GGT	TTT	AAA	TGA	CAG	CAT	8304
	Tyr	Leu	Ala	Ala	Asp	Pro	Cys	Arg	Ile	Asn	Gly	Phe	Lys	*	Gln	His	
					3325					3330					3335		

	GTT CTT TCA GTT TGA GCA AAA TTT GTG CAA TTG CAA AGA AGC TTT AGA	8352
	Val Leu Ser Val * Ala Lys Phe Val Gln Leu Gln Arg Ser Phe Arg	
	3340 3345 3350	
5	ATC ATG TGG AAC ATG CAC TTA CAT TTC ATC TGA CAA TAT AGG AAG GAG	8400
	Ile Met Trp Asn Met His Leu His Phe Ile * Gln Tyr Arg Lys Glu	
	3355 3360 3365	
10	AGC CCG ACG TCG CAT GCT CCT CTA GAC TCG AGG AAT TCG CAA GAT TGT	8448
	Ser Pro Thr Ser His Ala Pro Leu Asp Ser Arg Asn Ser Gln Asp Cys	
	3370 3375 3380	
15	CTG TCA AAA GAT TGA GGA AGA GGC AGA TGC GCA ATT TCT TTG TTT GTC	8496
	Leu Ser Lys Asp * Gly Arg Gly Arg Cys Ala Ile Ser Leu Phe Val	
	3385 3390 3395 3400	
20	TCA TGG TTT CTC AAG TAA GAC TTA TAT CTG ATC TCT TCA ATT TTT GAG	8544
	Ser Trp Phe Leu Lys * Asp Leu Tyr Leu Ile Ser Ser Ile Phe Glu	
	3405 3410 3415	
25	ATT GCC TGT TTT TCA CAA TGG CAT ATG TTG TCA GGT GAA ACA TCC AAT	8592
	Ile Ala Cys Phe Ser Gln Trp His Met Leu Ser Gly Glu Thr Ser Asn	
	3420 3425 3430	
30	CCC AGT ATT AAT AGA GCC AAC ATG AAG GGA TTG CTT ATC TGA GAT ATC	8640
	Pro Ser Ile Asn Arg Ala Asn Met Lys Gly Leu Leu Ile * Asp Ile	
	3435 3440 3445	
35	TGC CAA AGT TGA ATT CTT AGA TTC ACC TTC TTC AGT ATT TCA GAC CTT	8688
	Cys Gln Ser * Ile Leu Arg Phe Thr Phe Phe Ser Ile Ser Asp Leu	
	3450 3455 3460	
40	CTA AGC ATT TTC ATT TTT TTT TTC AAT TGT TAG GGA GTT CCA ATG TTT	8736
	Leu Ser Ile Phe Ile Phe Phe Phe Asn Cys * Gly Val Pro Met Phe	
	3465 3470 3475 3480	
45	TAC ATG GGC GAT GAA TAT GGC CAC ACA AAA GGG GGC AAC AAC AAT ACA	8784
	Tyr Met Gly Asp Glu Tyr Gly His Thr Lys Gly Gly Asn Asn Asn Thr	
	3485 3490 3495	
50	TAC TGC CAT GAT TCT TAT GTC AGT ACA ATT TGG TCA CAT ATT GTT GTT	8832
	Tyr Cys His Asp Ser Tyr Val Ser Thr Ile Trp Ser His Ile Val Val	
	3500 3505 3510	
55	CTA AGT AAC TAT CTT CAA ATC TTT GCA TTC ATC CGT CAT GGC TCT TCT	8880
	Leu Ser Asn Tyr Leu Gln Ile Phe Ala Phe Ile Arg His Gly Ser Ser	
	3515 3520 3525	
60	GTA GGT CAA TTA TTT TCG CTG GGA TAA AAA AGA ACA ATA CTC TGA CTT	8928
	Val Gly Gln Leu Phe Ser Leu Gly * Lys Arg Thr Ile Leu * Leu	
	3530 3535 3540	
65	GCA AAG ATT CTG CTG CCT CAT GAC CAA ATT CCG CAA GTA AGT ATT CCG	8976
	Ala Lys Ile Leu Leu Pro His Asp Gln Ile Pro Gln Val Ser Ile Pro	
	3545 3550 3555 3560	
70	TTG AAT AAT TTC TGT GTA GAA CCA CTG AAG GTG CCT CCA AAC GCT AAG	9024
	Leu Asn Asn Phe Cys Val Glu Pro Leu Lys Val Pro Pro Asn Ala Lys	
	3565 3570 3575	
75	CGA GCA AGG TCA ATT TCA CAC CCT AAT CAA GTT GGT GTT GTC TAT TTG	9072
	Arg Ala Arg Ser Ile Ser His Pro Asn Gln Val Gly Val Val Tyr Leu	
	3580 3585 3590	



- 115 -

	TGT ATT TGA TCT GCT GCA CTG TAG GGA GTG CGA GGG TCT TGG CCT TGA	9120
	Cys Ile * Ser Ala Ala Leu * Gly Val Arg Gly Ser Trp Pro *	
	3595 3600 3605	
5	GGA CTT TCC AAC GGC CGA ACG GCT GCA GTG GCA TGG TCA TCA GCC TGG	9168
	Gly Leu Ser Asn Gly Arg Thr Ala Ala Val Ala Trp Ser Ser Ala Trp	
	3610 3615 3620	
10	GAA GCC TGA TTG GTC TGA GAA TAG CCG ATT CGT TGC CTT TTC CAT GGT	9216
	Glu Ala * Leu Val * Glu * Pro Ile Arg Cys Leu Phe His Gly	
	3625 3630 3635 3640	
15	ACA CAT ATA GTT CTG ACA CTT CAC TAT AGT TGT TTT AAA AAA GAA AAT	9264
	Thr His Ile Val Leu Thr Leu His Tyr Ser Cys Phe Lys Lys Glu Asn	
	3645 3650 3655	
	TTA ACT CAA AAG TAA ATT ATG GAG A	9289
	Leu Thr Gln Lys * Ile Met Glu	
	3660	
20		

- 116 -

**CLAIMS**

1. A nucleic acid sequence encoding an enzyme of the starch biosynthetic pathway in a cereal plant, wherein the  
5 enzyme is selected from the group consisting of starch branching enzyme I, starch branching enzyme II, starch soluble synthase I, and debranching enzyme, with the proviso that the enzyme is not soluble starch synthase I of rice, or starch branching enzyme I of rice or maize.  
10
2. A sequence according to claim 1, wherein the sequence is a genomic DNA or cDNA sequence.
3. A sequence according to claim 1 or claim 2,  
15 wherein the sequence is functional in wheat.
4. A sequence according to any one of claims 1 to 3, wherein the sequence is derived from a *Triticum* species.
- 20 5. A sequence according to claim 4, wherein the *Triticum* species is *Triticum tauschii*.
6. A sequence according to any one of claims 1 to 5, wherein the sequence encodes starch branching enzyme I or a  
25 biologically-active fragment thereof, and wherein the sequence has at least 70% sequence homology with the sequence shown in SEQ ID NO:5 or SEQ ID NO:9.
7. A sequence according to claim 6, wherein the  
30 homology is at least 90%.
8. A sequence according to any one of claims 1 to 5, wherein the sequence encodes starch branching enzyme II a or biologically-active fragment thereof, and wherein the  
35 sequence has at least 70% sequence homology with the sequence shown in SEQ ID NO:10.

- 117 -

9. A sequence according to claim 8, wherein the homology is at least 90%.

10. A sequence according to any one of claims 1 to 5,  
5 wherein the sequence encodes soluble starch synthase or a biologically-active fragment thereof, and wherein the sequence has at least 70% sequence homology with the sequence shown in SEQ ID NO:11 or SEQ ID NO:13.

10 11. A sequence according to claim 10, wherein the homology is at least 90%.

12. A sequence according to claim 11, wherein the sequence encodes a 75 kD soluble starch synthase of wheat.

15 13. A sequence according to claim 12, which encodes an amino acid sequence at least 70% homologous to that shown in SEQ ID NO:14.

20 14. A sequence according to any one of claims 1 to 5, wherein the sequence encodes debranching enzyme or a biologically-active fragment thereof, and wherein the sequence has at least 70% sequence homology with the sequence shown in SEQ ID No:17.

25 15. A sequence according to claim 14, wherein the homology is at least 90%.

16. A promoter of an enzyme selected from the group  
30 consisting of starch branching enzyme I, starch branching enzyme II, starch soluble synthase I, and debranching enzyme, with the proviso that the enzyme is not soluble starch synthase I of rice, or starch branching enzyme I of rice or maize.

35 17. A promoter according to claim 16, wherein the promoter is a starch branching enzyme I promoter or

- 118 -

biologically-active fragment thereof, and wherein the promoter sequence has at least 70% sequence homology with the sequence shown in SEQ ID No:8.

5 18. A sequence according to claim 17, wherein the homology is at least 90%.

19. A promoter according to claim 16, wherein the promoter is a starch soluble synthase I promoter or  
10 biologically-active fragment thereof, and wherein the promoter sequence has at least 70% sequence homology with the sequence shown in SEQ ID No:15.

20. A sequence according to claim 19, wherein the  
15 homology is at least 90%.

21. A nucleic acid construct comprising a nucleic acid sequence encoding an enzyme of the starch biosynthetic pathway in a cereal plant, operably linked to one or more  
20 nucleic acid sequences facilitating expression of the nucleic acid sequence in a plant, wherein the enzyme is selected from the group consisting of starch branching enzyme I, starch branching enzyme II, starch soluble synthase I, and debranching enzyme, with the proviso that  
25 the enzyme is not soluble starch synthase I of rice, or starch branching enzyme I of rice or maize, a biologically-active fragment thereof.

22. A nucleic acid construct for targeting a gene to  
30 the endosperm of a cereal plant, comprising one or more promoter sequences selected from the group consisting of SBE I promoter, SBE II promoter, SSS I promoter, and DBE promoter, operatively linked to a nucleic acid sequence encoding a protein, wherein the expression of the targetted  
35 gene in the endosperm of a cereal plant is modified.

- 119 -

23. A construct according to either claim 21 or claim 22, wherein the promoter or nucleic acid sequence is also operatively linked to one or more additional targeting sequences and/or one or more 3' untranslated sequences.

5

24. A construct according to claim 23, wherein the nucleic acid encoding the protein is either in the sense or antisense orientation.

10 25. A construct according to claims 24, wherein the protein is an enzyme of the starch biosynthetic pathway.

26. A construct according to claim 25, wherein the nucleic acid encoding the protein is in the antisense  
15 orientation, and the enzyme is selected from the group consisting of GBSS, starch debranching enzyme, SBE II, low molecular weight glutenin, and grain softness protein I.

27. A construct according to claim 25, wherein the  
20 nucleic acid encoding the protein is in the sense orientation, and the enzyme is selected from the group consisting of bacterial isoamylase, bacterial glycogen synthase, and wheat high molecular weight glutenin Bx17.

28. A construct according to any one of claims 21 to  
25 27, wherein the plant is a cereal plant.

29. A construct according to claim 28, wherein the cereal plant is either wheat or barley.

30 30. A construct according to claim 29, wherein the cereal plant is wheat.

31. A construct according to any one of claims 21 to 30, wherein the construct is either a plasmid or a vector.

35

- 120 -

32. A construct according to claim 31, wherein the plasmid or vector is suitable for use in the transformation of a plant.
- 5 33. A construct according to claim 32, wherein the plasmid is selected from the group consisting of those depicted in Figures 22a to 22f.
34. A construct according to claim 32, wherein the  
10 vector is a bacterium of the genus *Agrobacterium*.
35. A construct according to claim 34, wherein the vector is *Agrobacterium tumefaciens*.
- 15 36. A method of modifying the characteristics of starch produced by a plant, comprising the steps of:
- (a) introducing a nucleic acid sequence encoding an enzyme of the starch biosynthetic pathway into a host plant, and/or
  - 20 (b) introducing an anti-sense nucleic acid sequence directed to a gene encoding an enzyme of the starch biosynthetic pathway into a host plant,
- wherein the enzyme is selected from the group consisting of starch branching enzyme I, starch branching  
25 enzyme II, starch soluble synthase I, and debranching enzyme, with the proviso that the enzyme is not soluble starch synthase I of rice, or starch branching enzyme I of rice or maize, and wherein if both steps (a) and (b) are used, the enzymes in the two steps are different.
- 30
37. A method according to claim 36, wherein the plant is a cereal plant.
38. A method according to claim 37, wherein the cereal  
35 plant is wheat or barley.

- 121 -

39. A method of targeting expression of a gene to the endosperm of a cereal plant, comprising the step of transforming the plant with a construct according to any one of claims 21 to 35.

5

40. A method of modulating the time of expression of a gene in endosperm of a cereal plant, comprising the step of transforming the plant with a construct according to any one of claims 21 to 35.

10

41. A method according to claim 40, wherein when expression at an early stage following anthesis is desired, the construct comprises either the SBE II, SSS I, or DBE promoter.

15

42. A method according to claim 40, wherein when expression at a later stage following anthesis is desired, the construct comprises the SBE I promoter.

20

43. A plant transformed with a construct according to any one of claims 21 to 35.

44. A plant according to claim 43, wherein the plant is a cereal plant.

25

45. A plant according to claim 44, wherein the cereal plant is wheat or barley.

30

46. A method of identifying variations in the starch synthesis characteristics of a cereal plant, comprising the step of identifying a variation in nucleic acid sequence in the intron regions of the SBE I, SBE II, SSS I or DBE genes.

35

47. A method of identifying variations in the starch synthesis characteristics of a cereal plant, comprising the step of identifying a variation in nucleic acid sequence compared to the sequence shown in one or more SEQ ID NO:5,

- 122 -

SEQ ID NO:7, SEQ ID NO:9, SEQ ID NO:10, SEQ ID NO:11, SEQ ID NO:13, SEQ ID NO:15, SEQ ID NO:16, or SEQ ID NO:17.

48. A method according to claim 47, in which a  
5 mutation or absence of a SBE I, SBE II, SSS I or DBE gene is detected.

49. A method according to either claim 47 or claim 48,  
in which the cereal plant is wheat or barley.

10 50. A product comprising plant material propagated  
from a plant transformed with a nucleic acid sequence  
encoding an enzyme of the starch biosynthetic pathway in a  
cereal plant, operably linked to one or more nucleic acid  
15 sequences facilitating expression of the nucleic acid  
sequence in a plant, wherein the enzyme is selected from the  
group consisting of starch branching enzyme I, starch  
branching enzyme II, starch soluble synthase I, and  
debranching enzyme, with the proviso that the enzyme is not  
soluble starch synthase I of rice, or starch branching  
20 enzyme I of rice or maize, a biologically-active fragment  
thereof.

51. A product comprising plant material propagated  
from a plant in which a gene was targeted to the endosperm  
of a cereal plant, by a nucleic acid construct comprising  
25 one or more promoter sequences selected from the group  
consisting of SBE I promoter, SBE II promoter, SSS I  
promoter, and DBE promoter, operatively linked to a nucleic  
acid sequence encoding a protein, wherein the expression of  
the targetted gene in the endosperm of a cereal plant is  
30 modified.

52. A product according to claim 50 or claim 51  
wherein the product is a food product.



1/44

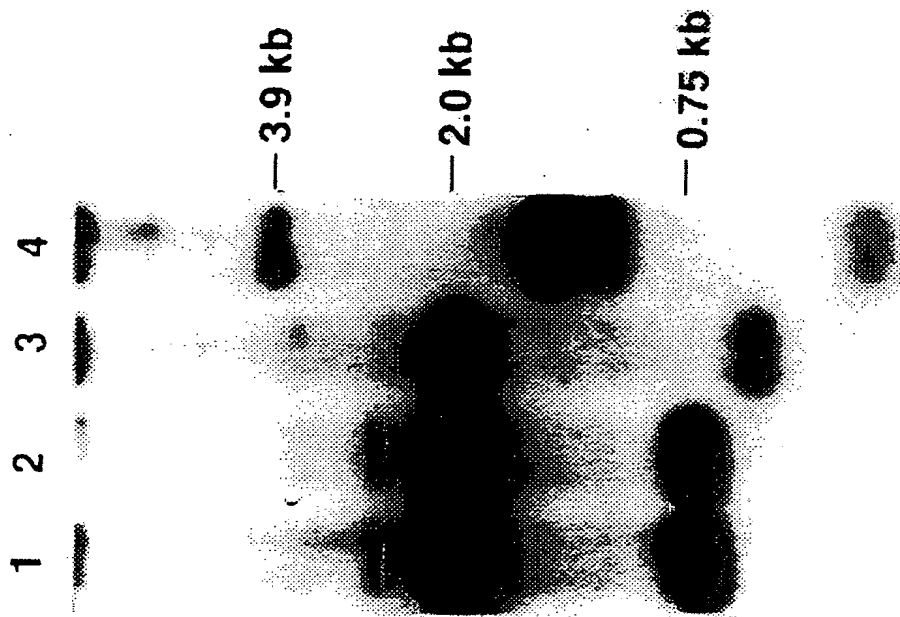


FIGURE 1

2/44

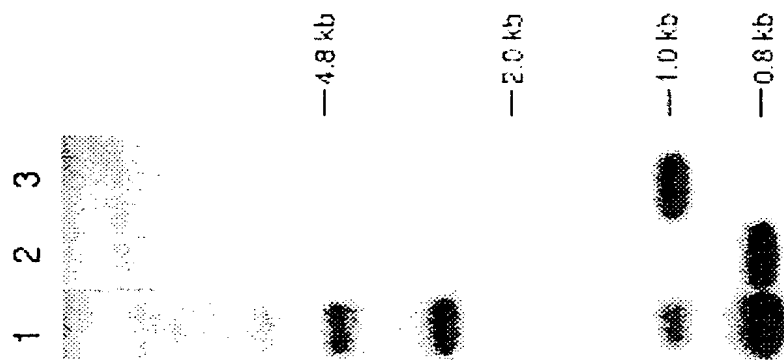
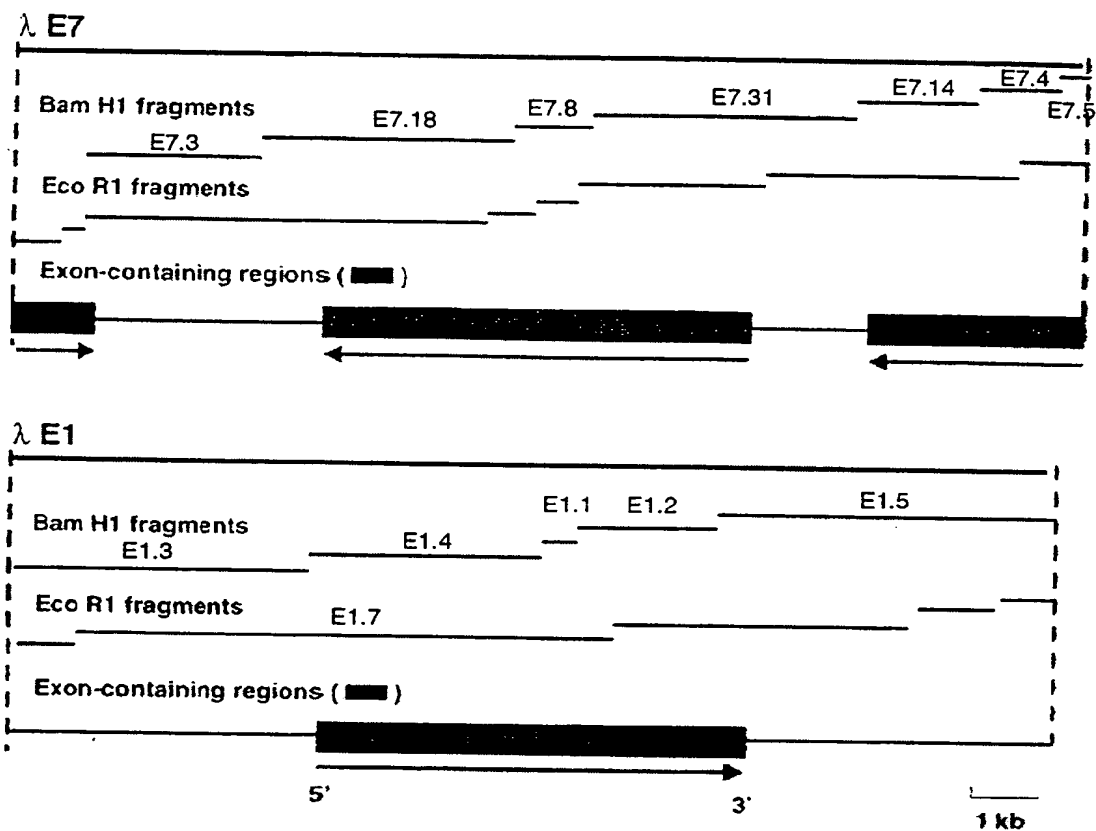


FIGURE 2

3/44



307716.1.1

FIGURE 3

## 4 / 44

	1				50
RSBEI	.....	*****	*....**pl	lp*****	**ag*****
MSBEI	.....	*****v*p**	**tplp***r	***h***aa*	pg*****
D4cDNA	.....	*****ap*c	**sl...**p	**pa***g*	**s*.....
PESBEII	.....	.....	.....	.....	.....
POSBE	meinfkvlsk	pirgsfp*f*	pkv*sgas*n	kic*psqh*t	*lkf*squers
D2cDNA	.....	*****s*ll	prp*a*....	....*****l*	*****ggk
Consensus	-----	-MLCLTSSSS	SP-S-APPR-	SRS-ADRPSP	GIIAGGGNVR
	51				100
RSBEI	l..**v*... *	p*****g**	*tn***pa**	rk*****v*v	***..*****
MSBEI	l..**l**qc	ka***gv***	*****ataa*v	q*d*****ak	g**..*****
D4cDNA	.....	.....	*****p*s*	prdy*****a	*g*..gd***
PESBEII	.....	.....	.....mt	d*ks**psv*	*f*..nig*
POSBE	w..d*s*t*k	*rv*kde*mk	h*saisa*lt	d**s***pl*	***kt*nigl
D2cDNA	rlsv*p***f	ll**l****a	***sf*s***	rg**ia***.	tgygs*****
Consensus	---SV-SVP-	S-RRSWPRKV	KSKFSV-VTA	-DNKTMAT-E	EDV--DHLPI
	101				150
RSBEI	*****e*	*****n**i**	*****c****	*****v	*****v
MSBEI	*****i*	*****s****	*****gs**e	n**s**s***	*****n
D4cDNA	*****ag*	*****s****k	*****s****	*****s****	*****s****
PESBEII	lnv**ss**p*	*****k*****	**h**k***e	y*****q**a*	*****f*r*
POSBE	ln***t**p*	l*****h*****	*v*****m****	y**p*****aq	*****f*r*
D2cDNA	****l**ae*	*****d*trn*	*i*****s****	****s*****s****	*****s****
Consensus	YDLDPKLE-F	KDHFYRMRKR	YLDQKHLIEK	HEGGLEEFKSK	GYLKFGINTE
	151				200
RSBEI	*g*****	*****s****	*****ak*	*****k****	**k*****
MSBEI	*dg*****	*****e****	***d***a**	*****k****	**k*d**k**
D4cDNA	nd*****	***m*****s****	*****g*	r*t**n****	*****s****
PESBEII	*dgis*****	*****i****	***g*****l	h*****q****	**q*pdad*n
POSBE	*gci*****	*****dev**	***g*****s****	m*****q****	*****pd*ds*
D2cDNA	hg*s*****	***e*****s****	*****g*	**a**n****	*****s****
Consensus	--ATVYREWA	PAAQEAQLIG	DFNNWNGSNH	KMEKD-FGVW	SIRISHVNGK
	201				250
RSBEI	*****	***r**g*a*	*****s****	**f*****	*****s****
MSBEI	*****	***l*.g***	*****l***	*****s****	*****s****
D4cDNA	*****	***hr*d*l*	*****s****	**f*****	*****s****
PESBEII	*****r**	***k*sd***	*****k*	***ptr*a*	*****y****
POSBE	*v*****r**	***k**n***	*****k*	**a**t**a*	*****y****
D2cDNA	*****	***r*.h***	**q*****s****	***t**es**	****l*****
Consensus	PAIPHNSKVK	FRF-HG-GVW	VDRIPAWIRY	ATVDASKFGA	PYDGVHWDPP
	251				300
RSBEI	ac*****	*****s****	*****s****	*****s****	*****s****
MSBEI	a****t****	**s**a****	*****s****	k*a*****	*****s****
D4cDNA	sg*****	**r*****s****	*****s****	r*****s****	*****s****
PESBEII	l****q****	*****k****	*****ss	**r*ns****	**d*****e
POSBE	p****h*y*	*****r****	*****ss	**r*ns****	**d*****k*
D2cDNA	s*****n*	*****v****	*****v**g	kl*ag*****	p*****cl**
Consensus	-SERYVFKHP	RPPKPDAPRI	YEAHVGMSE	EPEVSTYREF	ADNVLPRIIRA

Figure 4

5/44

	301				350
RSBEI	*****	*****	*****	*****	*****
MSBEI	*****	*****	*****	*****	*****
D4cDNA	*****	*****ilcf*	w*****	*****	*****
PESBEII	*****	*****	w*****kp**	*****s**	*****
POSBE	*****	*****g**	*****	*****y*n**	*****
D2cDNA	t*****g	*****ds**	*****	*****	*****
Consensus	NNYNTVQLMA	IMEHSYYASF	GYHVTN-FFA	VSSRSGTPED	LKYL-DKAHS
	351				400
RSBEI	*****	*****	*****n	*h*****t**	*****
MSBEI	*****	*****	*****	*****a**	*****
D4cDNA	*****	*****s*m**	*****n	*****t**	*****
PESBEII	**n*****	*****	*****	s*q*****a**	*****
POSBE	**q**v***	*****	*****g	s*****a**	*****
D2cDNA	*****	*****i*	*****	ah*****yt**	k**n***ng*
Consensus	LGLRVLMDVV	HSHASNNVTD	GLNGYDVGQS	TQESYFH-GD	RGYHKLWDSR
	401				450
RSBEI	*****	*****	*****	*****	*****k*****
MSBEI	*****	*****	*****	*****	*****v*****
D4cDNA	*****	*****	*****	*****n	*****s*a*
PESBEII	*****ks.	s*****	*****k*****	*****	*****a****
POSBE	*****	*****	*****n*****	*****v	*****
D2cDNA	*****	*****	*****	*v*****n	*n*****s*n*
Consensus	LFNYANWEVL	RFLLSNLRW	-DEFMFDGFR	FDGVTSMLYH	HHGINMGFTG
	451				500
RSBEI	*****	*****	*****l**	*****	*****
MSBEI	**q*****	a*****	*****l**	*****	*****
D4cDNA	*****g***	*****	*****i**	*****	*****s**
PESBEII	d*n*****e**	*****	**s*v*di**	***d*****	***g*g***s
POSBE	**n*****ea*	*****	**n*i**i**	*****	***g*g***s
D2cDNA	*****ig***	n***f*****	*****l**	**i***v***	*****
Consensus	NYKEYFSLDT	DVDAVVYML	ANHLMHK-LP	EATVVAEDVS	GMPVLCRPVD
	501				550
RSBEI	*****	*****	*****rk*	****.*vq**	*****
MSBEI	*****	*****	*****	**g*.*ah**	*****
D4cDNA	*****	*****	*****l**	***a.*ah**	*****
PESBEII	*v*****	*****k***	*****k***	**k*.*sln*	*****
POSBE	*****	*****k***	*****n*e**	**k*.*tss*	*****
D2cDNA	***l*****q	**t*****	**e*g*qq*	**sv*sq**	*****p**f*
Consensus	EGGVGFDYRL	AMAIPDRWID	YLKNKDDSEW	SMSE-I--TL	TNRRYTEKCI
	551				600
RSBEI	*****	*****	*****t**	*****n	*****
MSBEI	*****	*****	*****t**	*****	*****
D4cDNA	*****	*****m****	*****t**	*****	*****
PESBEII	s*****	*****	**e***ss*	c*tml*****	***s*h****
POSBE	*****	*****	*****s***	c*td***v**	*****h****
D2cDNA	****rqnh**	**s*m****	**w*t*s***	a*d*d*****	*a*****
Consensus	AYAESHQSI	VGDKTIAFLL	MDKEMY-GMS	DLQPASPTID	RGIALQKMIH

Figure 4 (cont..)

6/44

	601				650
RSBEI	*****	*****	*****	*****	*****
MSBEI	*****	*****	*****	*****	*****
D4cDNA	*****	*****	*****	*****	*****s i-
PESBEII	*****	*****	*****	**g*****	lt**n**n
POSBE	*f*****	*****	*****	*****	****n*a*s*
D2cDNA	*****s	**k*****	.....	.....	.....
Consensus	FITMALGGDG	YLNFMGNEFG	HPEWIDFPRE	GNNWSYDKCR	-RQWSLVDTD
	651				700
RSBEI	*****	*****e	*****	*****k***	*****
MSBEI	*****	*****r	*****	*****	*****
D4cDNA	*****	*****	*****	*****k**	*****
PESBEII	*****	*r***l***	**i*a*t***	**st*n***	*****
POSBE	*****	*r***s***	***a*g***	**s*d**n**	*****
D2cDNA	.....*****	v**vdtps**	c*****n*t	a*h*****g	sa*tk*....
Consensus	HLRYKYMNAF	DQAMNALD-K	FSFLSSSKQI	VSDMNEE-KV	IVFERGDLVF
	701				750
RSBEI	*****n***	k*****	*****	**v*****	*****
MSBEI	*****k***	*****	*****	**v*****	*****
D4cDNA	*****s***	*****	***k*****	**m*****	aqyn*****
PESBEII	*****en**	*****	*****	*te*****	***a*q****
POSBE	*****kn**	*****	*****	*we*****t	*****
D2cDNA	.*thlrsgc*	*p.....s**	stssc**...	.*gpsnqspf	skpfig*pgc
Consensus	VFNFHP-KTY	EGYKVGCDLP	GKYRVALDSD	AL-FGGHGRV	GHDVDHFTSP
	751				800
RSBEI	**m*****	*****	.....	.....	*****
MSBEI	*****	*****	.....	.....	*****
D4cDNA	*****	*****	.....	.....	*****
PESBEII	*****	*****	.....	.....	*****h**v*
POSBE	*****	**g*qipskc	c llrehvwli	telmnacq* l	kitrq*f*vs
D2cDNA	ifcc*lfkge	*.....	.....	.....	.....
Consensus	EG-PGVPETN	FNNRP-----	-----	-----NSFKV	LSPPRTCVAY
	801				850
RSBEI	*...****dr	**l*rg**va	s**i.vte**	**e**s....	..**ti**gw
MSBEI	*...****ag	agr*lhak*e	t***s**es*	**k*s*....	..a....ssk
D4cDNA	*...****ka	*kpkde****	w**aa*g.*	**e***vkda	ad**at**sk
PESBEII	*...****q	**snpnlg*	*ee**a*adt	**aripdvs*	e*..ed*nld
POSBE	*yqqp*sr*v	trnlkiry lq	*sv**tna*q	klkf**qtf*	v*yyqqp ilr
D2cDNA	.....	.....	.....	.....	.....
Consensus	Y---RVDER-	EE-R--GAAS	-GKT-PA-YI	DV-ATR----	-SGE--SG--
	851				876
RSBEI	kg***d*cg*	**mk***r**	*e*c*d		
MSBEI	edk*atagg*	**wk*arqp*	*q*t**		
D4cDNA	ka*ttgg*ss*	**in***g*p	*k*n*		
PESBEII	r*e*ns**av	dagi*kvere	vvgdn*		
POSBE	r*tr*lk*sl	stnist*...	.....		
D2cDNA	.....	.....	.....		
Consensus	--SEK-DD-K	KG--FVF-SS	D-D-K-		

Figure 4 (cont..)

7/44

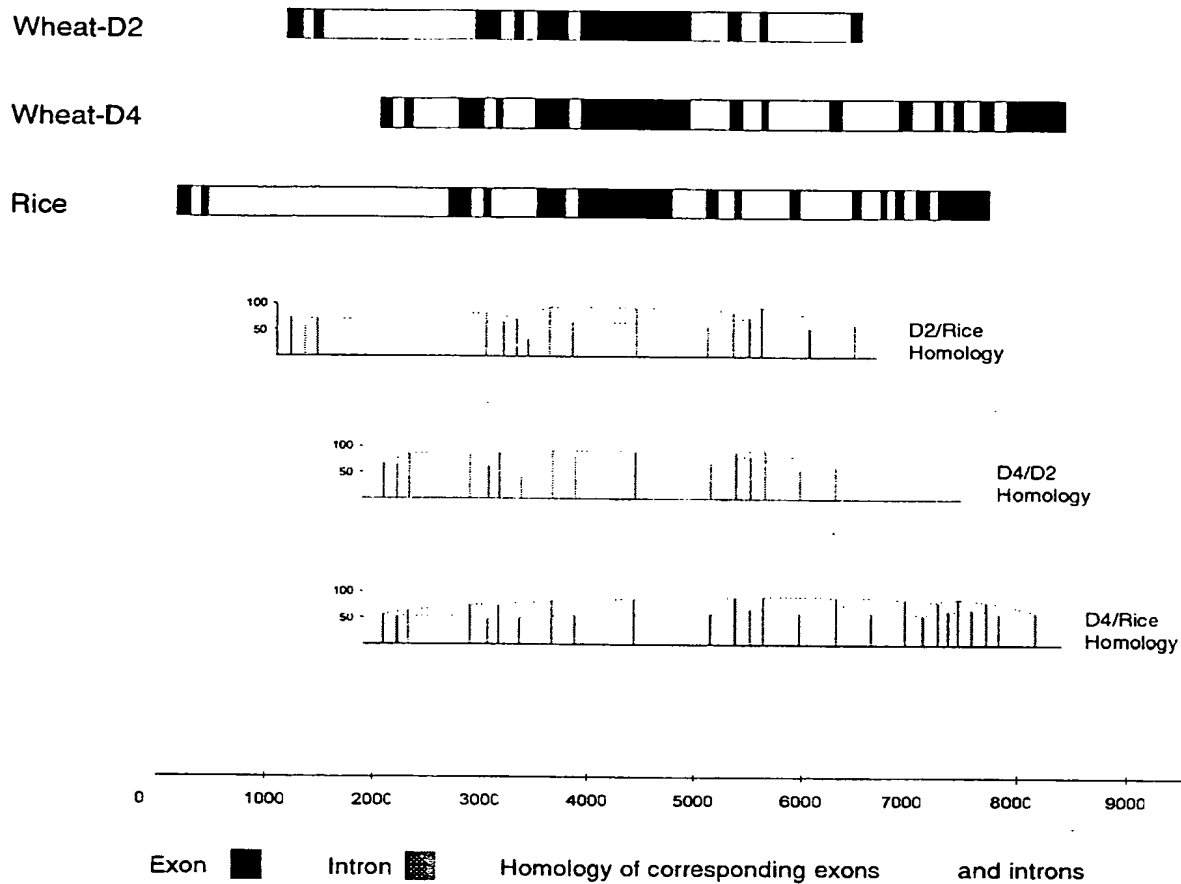


FIGURE 5

8/44

5' TCCCGTGTCTGCGCCAAGAGACTACACCATGGCAACAGCTGAAGATGGTGTGGCGACCT 5'  
3' AGGGCACAGACGCGGTTCTCTGATGTGGTACCGTTGTCGACTTCTACCAACCGCTGGA 3'

DNA

[ S R V C A K R L H H G N S \* R W C W R P  
P V S A P R D Y T M A T A E D G V G D L  
P C L R Q E T T P W Q Q L K M V L A T F

possible  
reading  
frames

[ V S A P R D Y T M A T A E D G V  
true N-  
terminal  
sequence  
for BE-1  
(Morell et  
al, 1997)

Figure 6



9/44

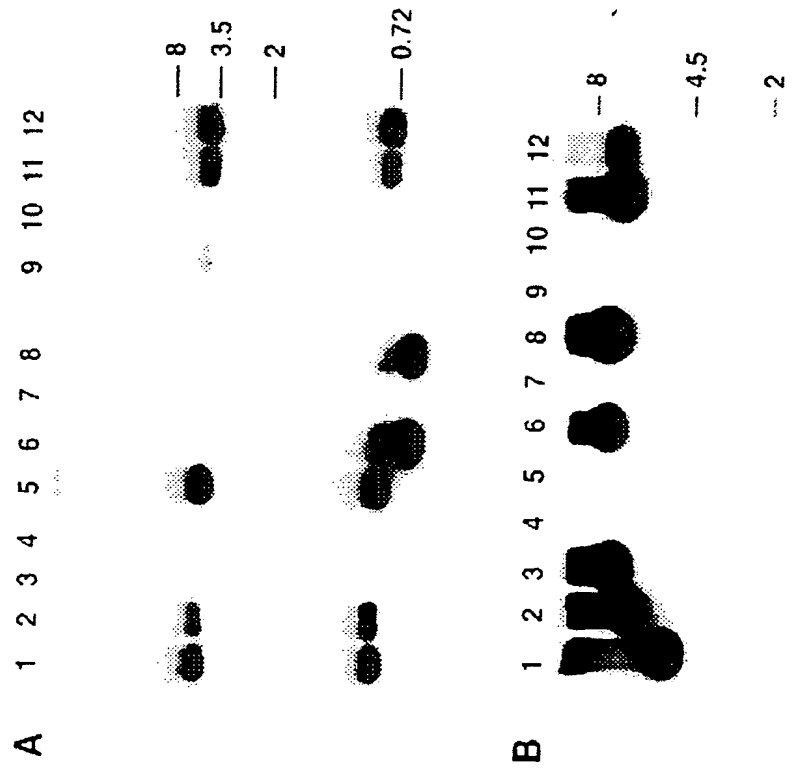


FIGURE 7

10/44

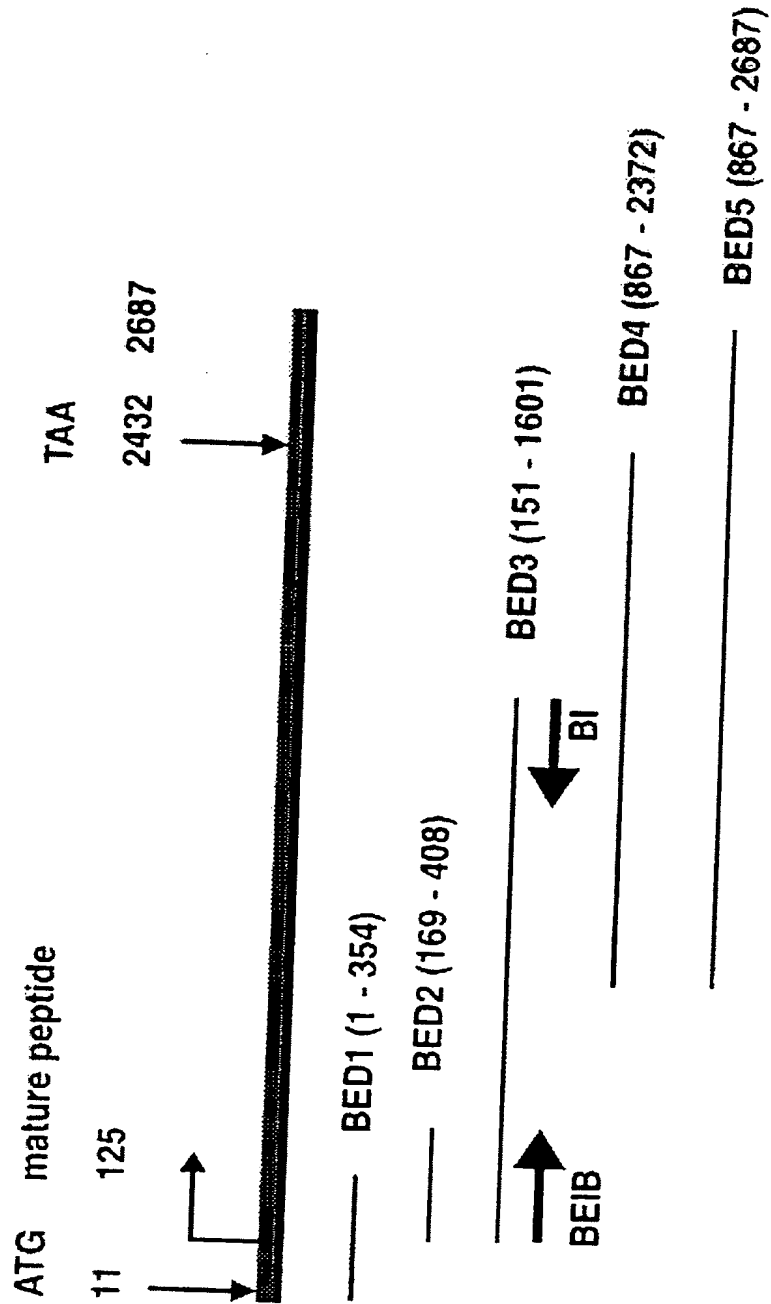


FIGURE 8

11/44

## Expression of Starch Biosynthetic Genes

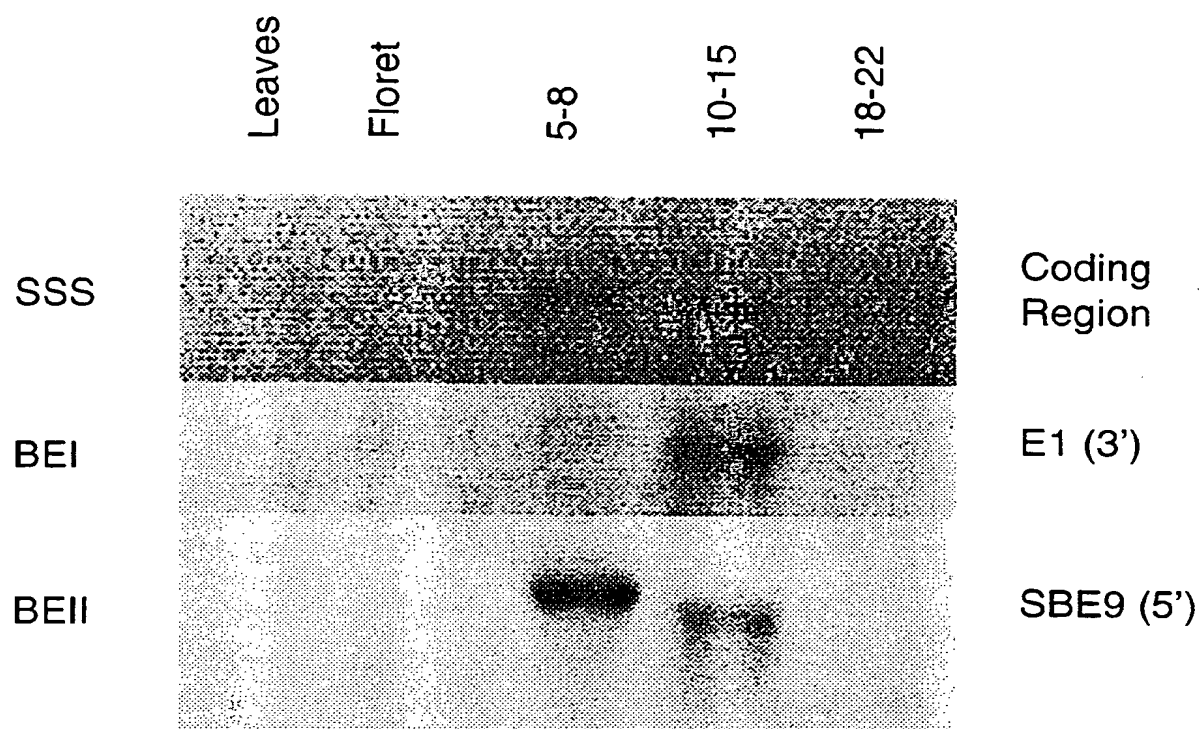


FIGURE 9A

12/44

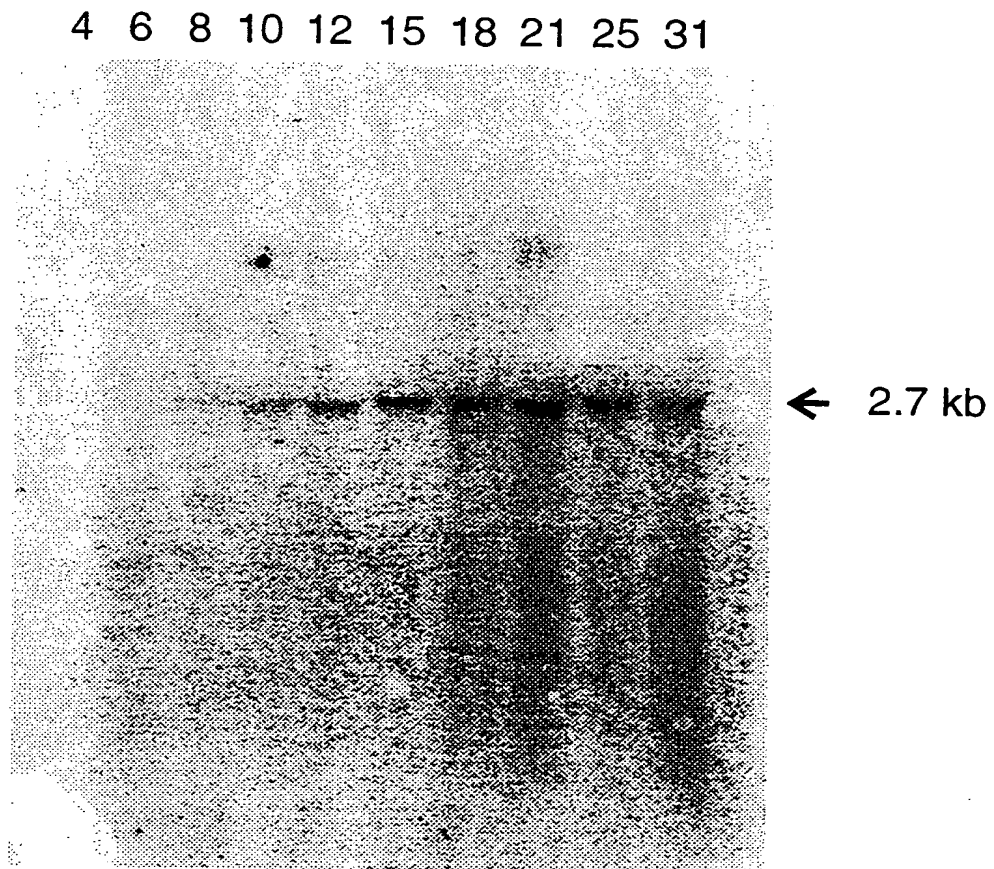
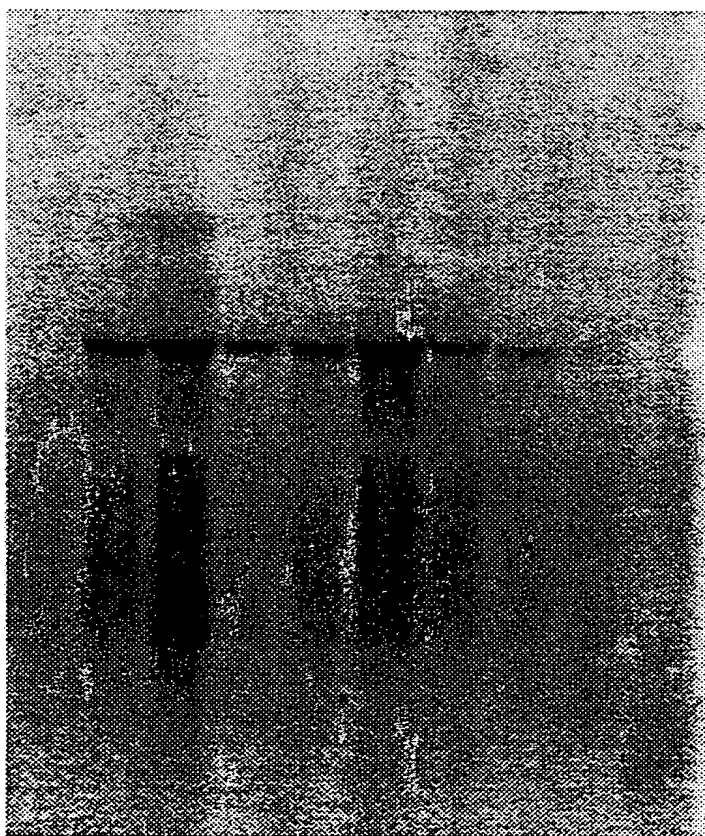


FIGURE 9B

13/44

4 6 8 10 12 15 18 21 25 31



← 2.9 kb

FIGURE 9C

14/44

4 6 8 10 12 15 18 21 25

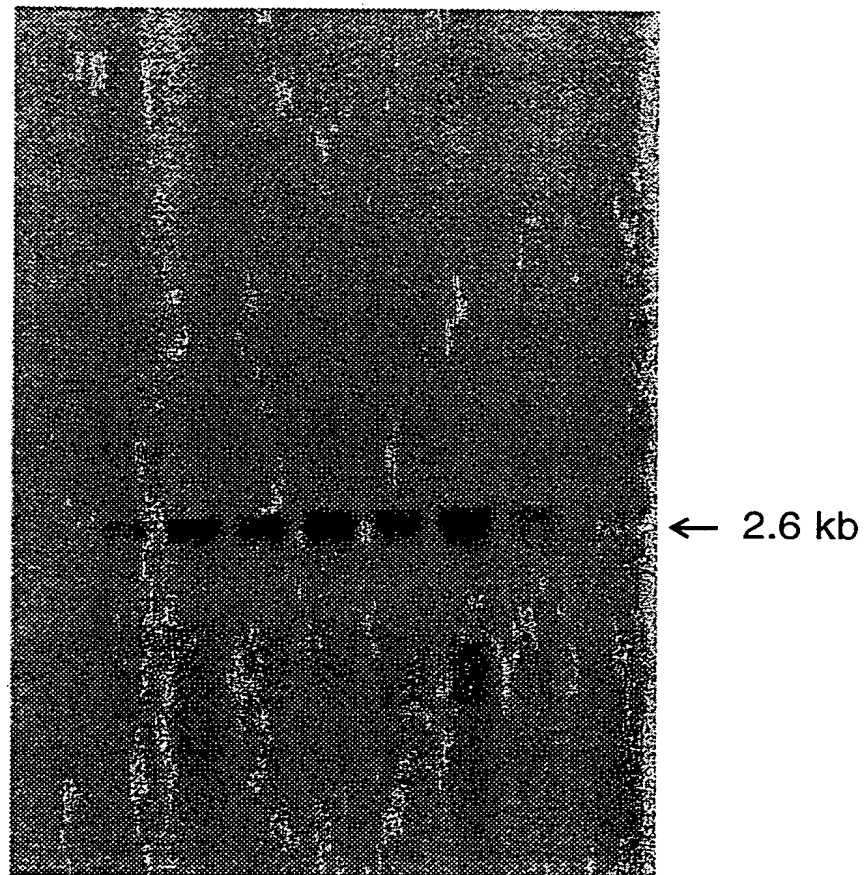


FIGURE 9D

15/44

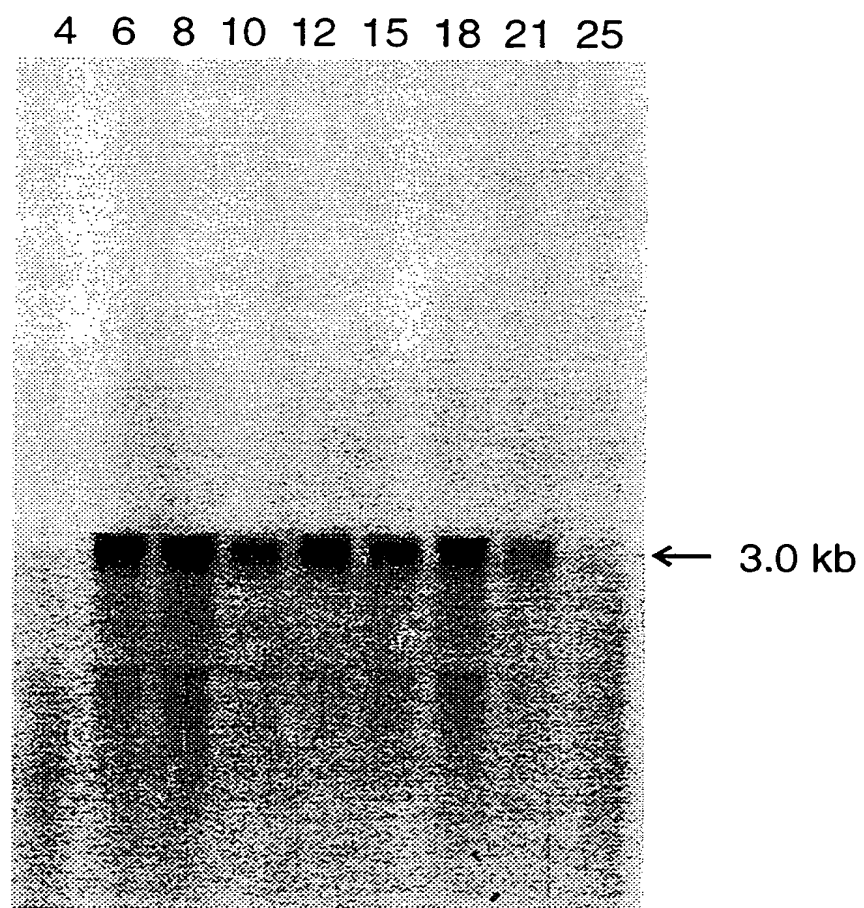


FIGURE 9E

16/44

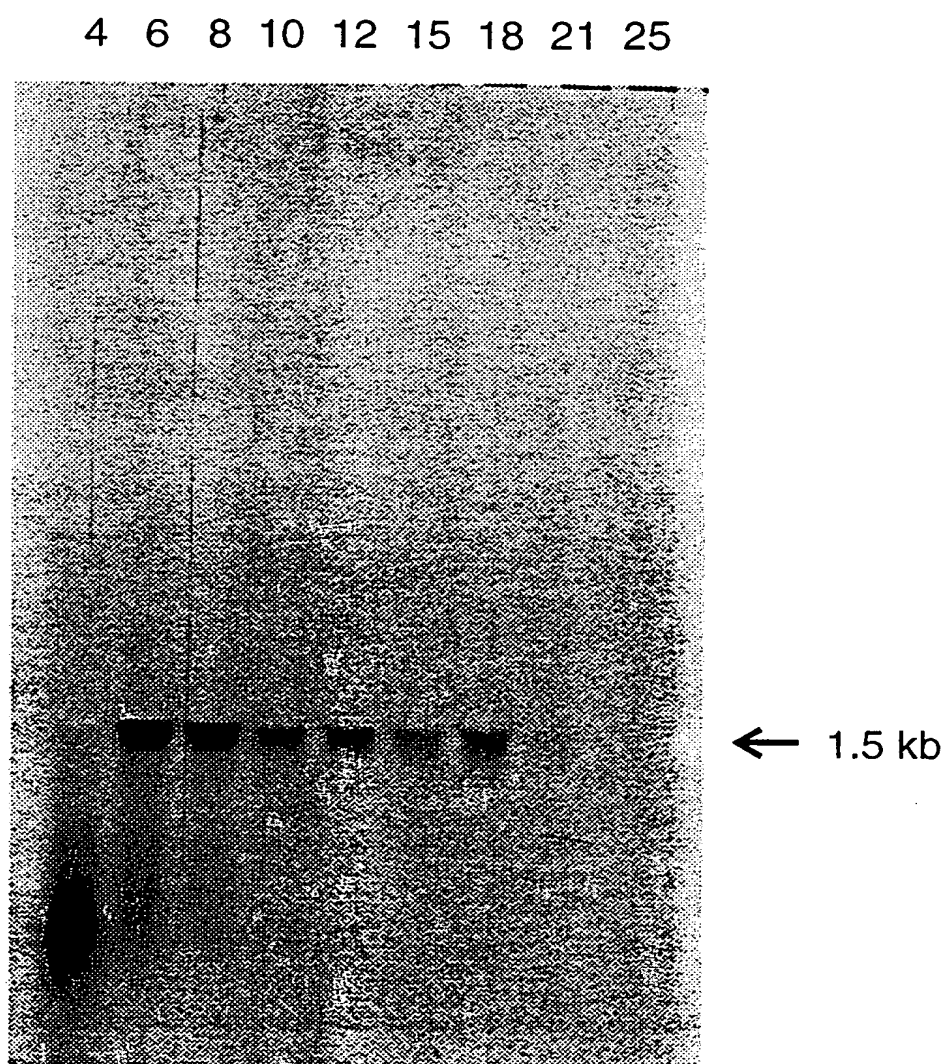


FIGURE 9F



17 / 44

4 6 8 10 12 15 18 21 25

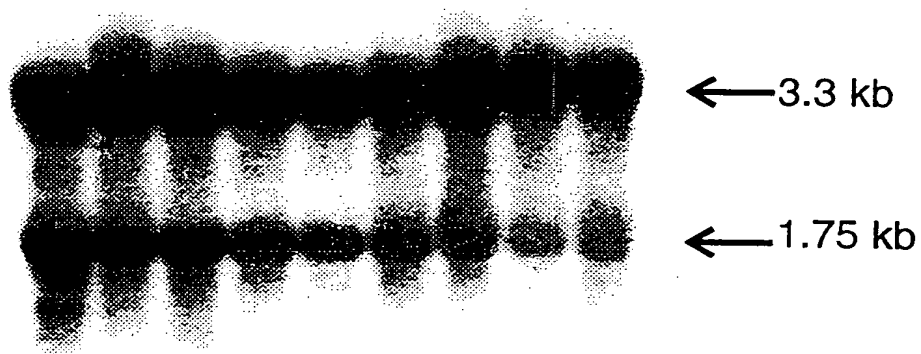


FIGURE 9G

18/44

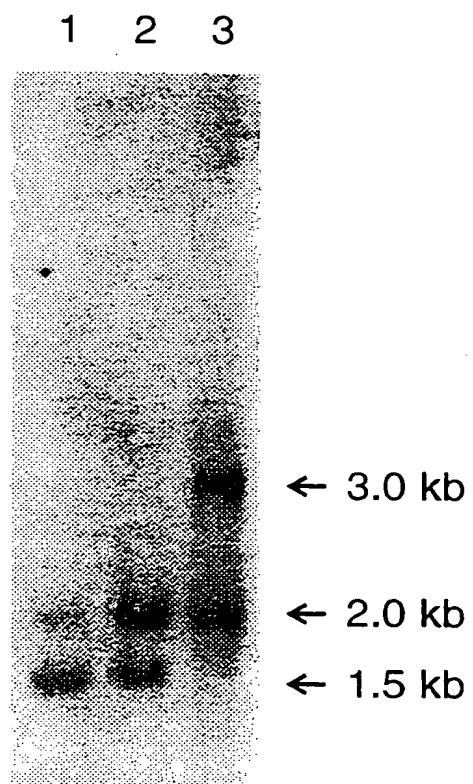


FIGURE 9H

19/44

DOTPLOT of: d10838.pnt Density: 12614.77 February 18, 1997 11:43

COMPARE Window: 21 Stringency: 14.0 Points: 20,788

sr427.res ck: 6,362, 1 to 11,099

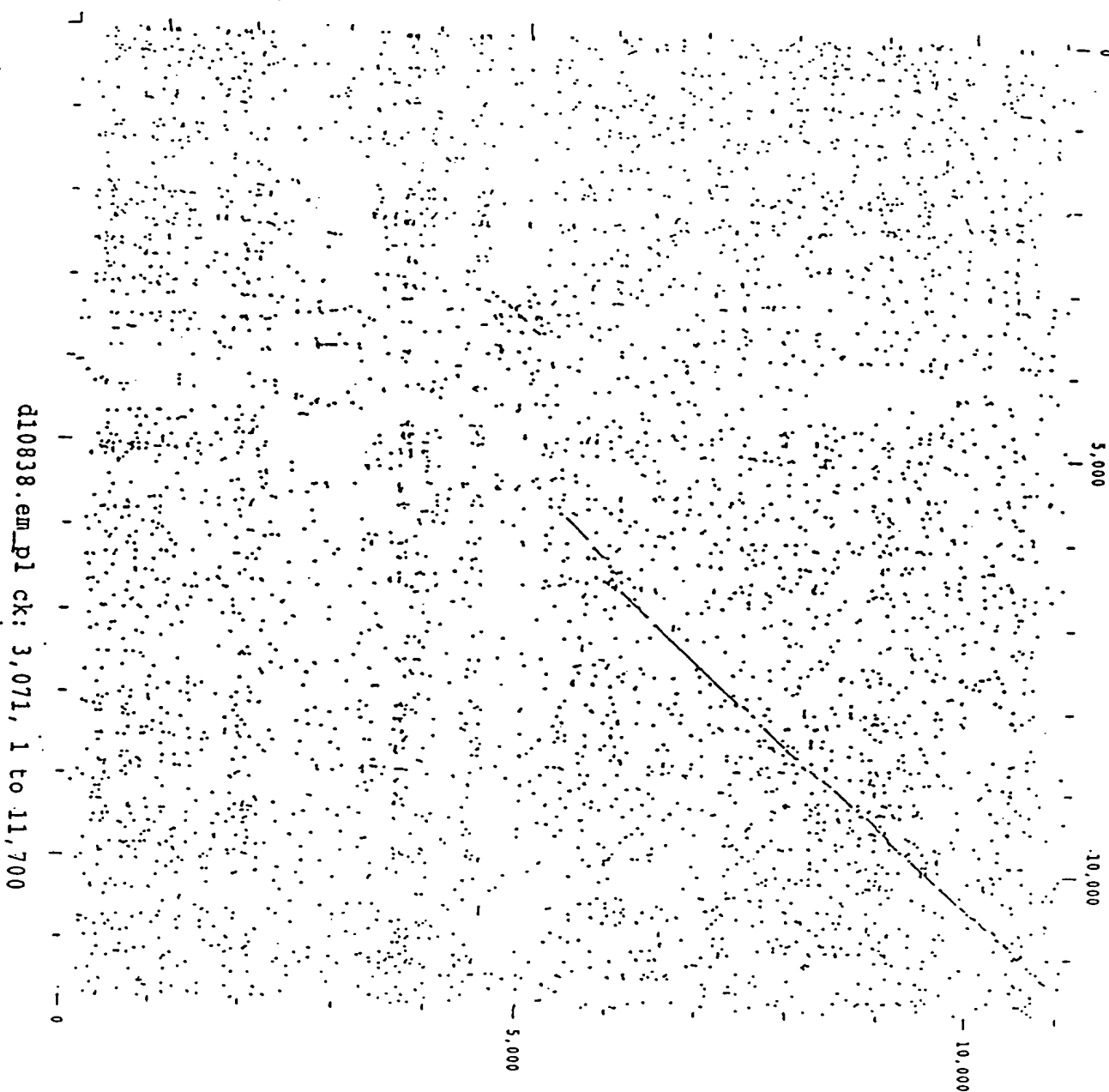


Figure 10

20/44

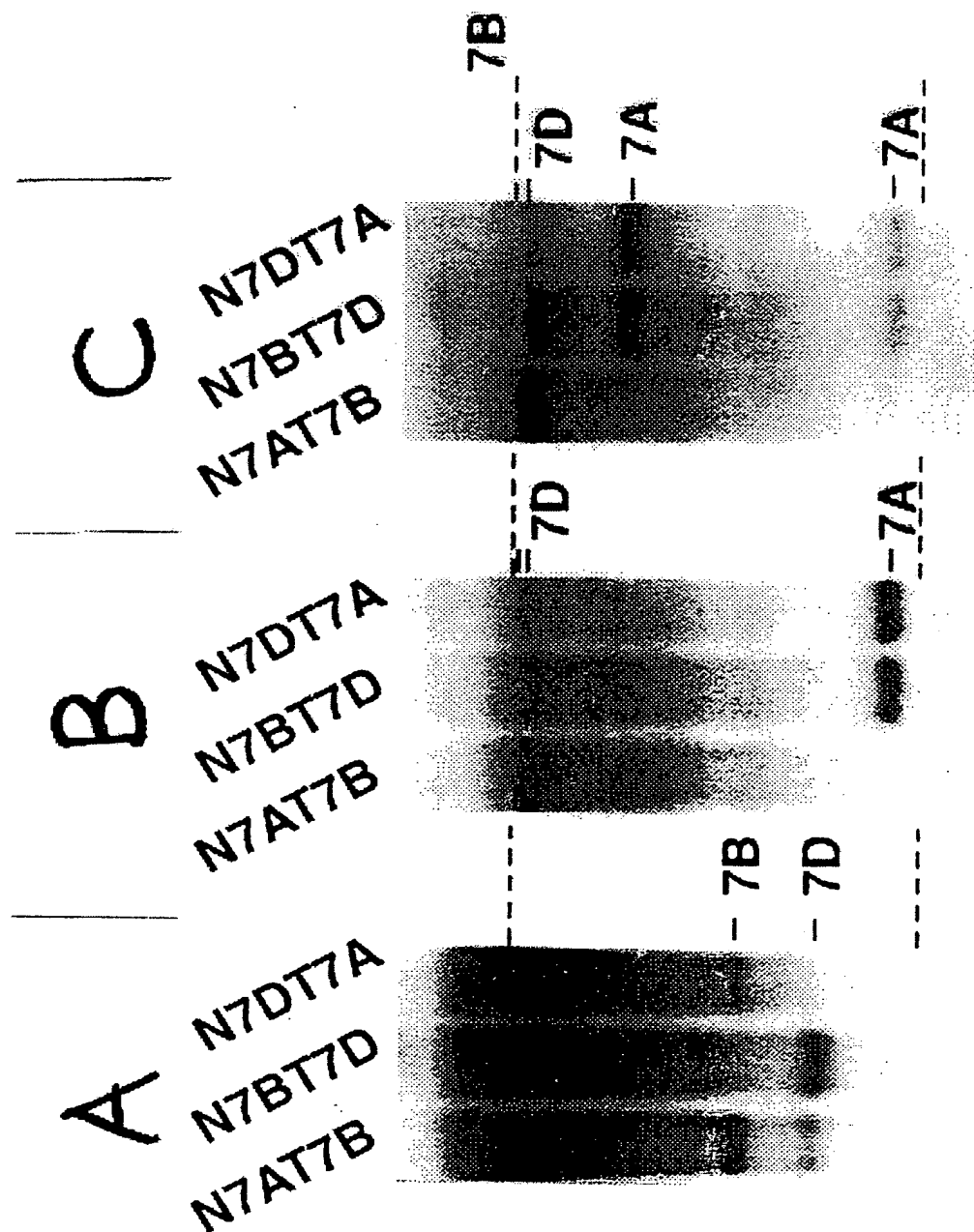


FIGURE 11

21/44

Genomic Clones from *T. tauschii*  
for SBE II.

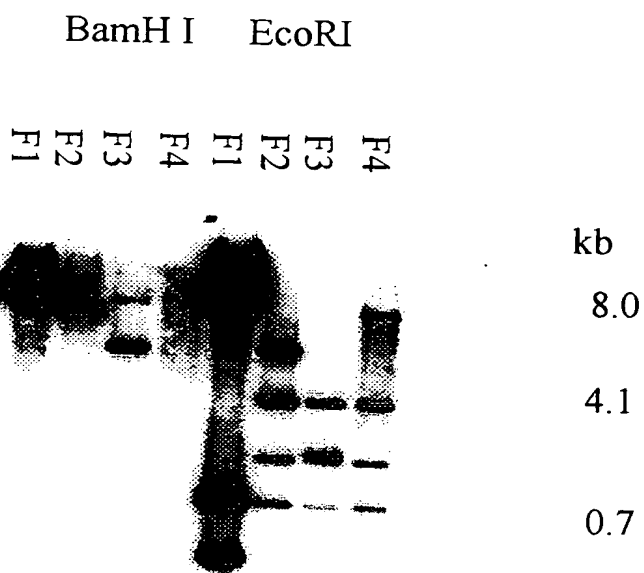


FIGURE 12

N-terminal sequences of cereal starch branching enzymes

Protein	1	2	3	4	5	6	7	8	9	1	1	1	1	1	1	1	1	2	2	2
<sup>a</sup>									0	1	2	3	4	5	6	7	8	9	0	1
RICEBEI <sup>b</sup>	A	T	A	R	K	N	K	T	M	V	T	V	V	E	E	V				
WBE-I <sub>AD</sub>	V	S	A	P	R	D	Y	T	M	A	T	A	E	D	G	V				
MAIZE	A	T	V	Q	E	D	K	T	M	A	T	A	K	G	D	V				
BEI <sup>c</sup>																				
RICEBEII <sup>d</sup>	A	A	G	A	S	G	E	-	V	M	I	P	E	G	E	S	D	G	M	P
WBE-II	A	A	S	P	G	K	-	V	L	V	P	D	G	E	S	D	D	L	A	S
MAIZE	A	A	A	A	R	K	A	V	M	V	P	E	G	E	N	D	G	L	A	S
BEI <sup>e</sup>																				

<sup>a</sup> N-terminal amino acid of the mature polypeptide. <sup>b</sup> Kawasaki *et al.*(1993), <sup>c</sup> Baba *et al.* (1991),

<sup>d</sup> Mizuno *et al.* (1993), <sup>e</sup> Fisher *et al.* (1993)

Residues in the wheat sequences showing identity with the respective maize or rice branching enzyme isoforms are highlighted in bold text.

Figure 13a

```

TTCCTTTTCTTTCTTTGGGNGGCGATGGCCGTTGCGATGNTGTTCCCAATGAATTT
1-----+-----+-----+-----+-----+ 60
AAGGGAAAAAAGAAACCCNCCCCCTACCGGACAACCTACNACAAGGGGTTACTTAAA

a   F P F F F F G ? G M A C W M ? F P N E F -
b   S L F F S L G G G W P V G ? C S P M N F -
c   P F F F L W ? G D G L L D ? V P Q * I S -

CCATGGAGTGAGAGAGATAGTTGGATNAGGGATCGCGNTTCCNGGAACGTGTATTTTTC
61-----+-----+-----+-----+-----+ 120
GGTACCTCACTCTCTCTATCAACCTANTCCCTAGCGCNAAGGNCCTTGACATAAAAAAAG

a   P W S E R D S W ? R D R ? S ? N C I F F -
b   H G V R E I V G ? G I A ? P G T V F F S -
c   M E * E R * L D ? G S R F ? E L Y F F P -

CCCNGCGGGGAAATGGCGTTAGTGTNACCCAGGCCCCTGGTGTACCACGGCTTTGATC
121-----+-----+-----+-----+-----+ 180
GGGNCGCCCCCTTTACCGCAATCACAGNTGGGTCCGGGACCACAATGGTGCCGAAACTAG

a   P ? G G N G V S V ? P G P G V T T A L I -
b   P A G E M A L V S T Q A L V L P R L * S -
c   ? R G K W R * C ? P R P W C Y H G F D H -

ATCTTCGTTTCATTCTGATATATATTTCTCATTCTTTTCTTCTGTTCTTCTGCTGIAA
181-----+-----+-----+-----+-----+ 240
TAAGAAGCAAAGTAAGACTATATATAAAAGAGTAAGAAAAAGAAGGACAAGAACGACATT

a   I L R F I L I Y I F S F F F F L F L L * -
b   F F V S F * Y I F S H S F S S C S C C N -
c   S S F H S D I Y F L I L F L P V L A V T -

CTGCAAGTGTGGGGTTTTTTCACATTTGTAGTCATCCTTGCATTTTGCAGGCGCGTCC
241-----+-----+-----+-----+-----+ 300
GAOGTTCAACACCGCAAAAAAGTGATAACATCAGTAGGAACGTAAACGTCCGCGGCAGG

a   L Q V V A F F H Y C S H P C I L Q A P S -
b   C K L W R F F T I V V I L A F C R R R P -
c   A S C G V F S L L * S S L H F A G A V L -

TGAGCGCGCGGCTCTCCAGGGAAGGTCTGCTGCGCTGACGGGAGAGNGAOGACTTGG
301-----+-----+-----+-----+-----+ 360
ACTCGGCGCGCGGAGAGGTCCCTTCCAGGACCACGGACTGCCGCTCTCNCCTGCTGAACC

a   * A A R P L Q G R S W C L T A R ? T T W -
b   E P R G L S R E G P G A * R R E ? R L G -
c   S R A A S P G K V L V P D G E ? D D L A -

CAAGTCCGGCGCAACCTGAAGAATTACAGGTACACACACTCGTGCCGTTAAATCTTCATA
361-----+-----+-----+-----+-----+ 420
GTTCAGGCGCGGTTGGACTTCTTAATGTCCATGTGTGTGAGCACGGCCATTTAGAAGTAT

a   Q V R R N L K N Y R Y T H S C R * I F I -
b   K S G A T * R I T G T H T R A G K S S Y -
c   S P A Q P E E L Q V H T L V P V N L H T -

CAATCGTTATTCACTTACCAAAATGCCGGATGAAACCAACCACGGATGCGTCAGGTTTCCA
421-----+-----+-----+-----+-----+ 480
GTTAGCAATAAGTGAATGGTTTACGGCCTACTTTGGTTGGTGCTACGCAGTCCAAAGCT

a   Q S I F T Y Q M P D E T N H G C V R F R -
b   N R Y S L T K C R M K P T T D A S G F E -

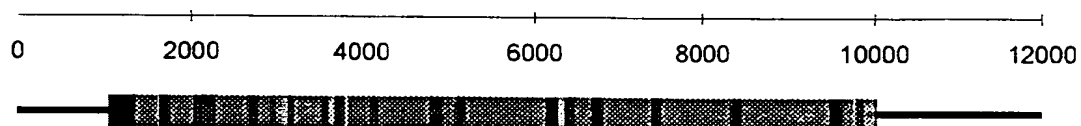
```

Figure 13b

24 / 44

*Branching Enzyme-II Genes*

Intron/Exon structure of wheat BE-II



Schematic Diagram of a cDNA for BE-II

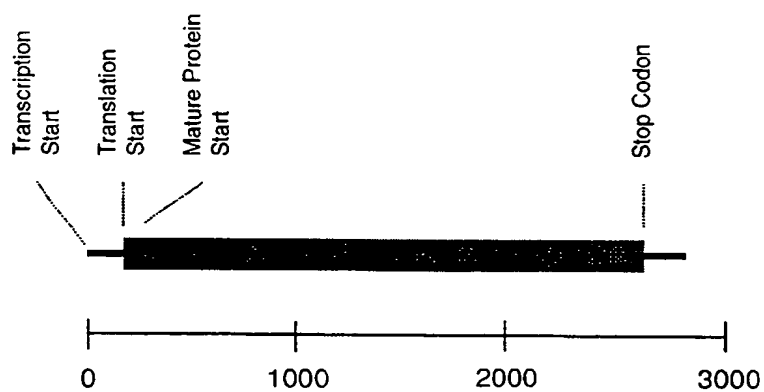


FIGURE 14



25/44

Wheat DNA probed with the  
5' conserved sequence of SBE II.

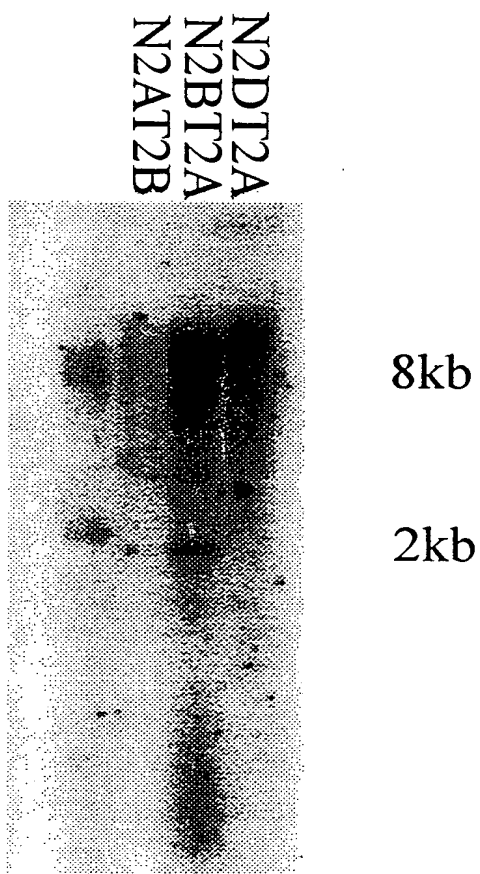


FIGURE 15

26/44

# COMPARISON OF N-TERMINAL SEQUENCES OF SOLUBLE STARCH SYNTHASE

GRYVAELSRFGPAARP      Deduced from wheat cDNA

GPYVAELSPFGPAAPP      Wheat N-terminal

Figure 16

27/44

## Soluble Starch Synthase Genomic Clones

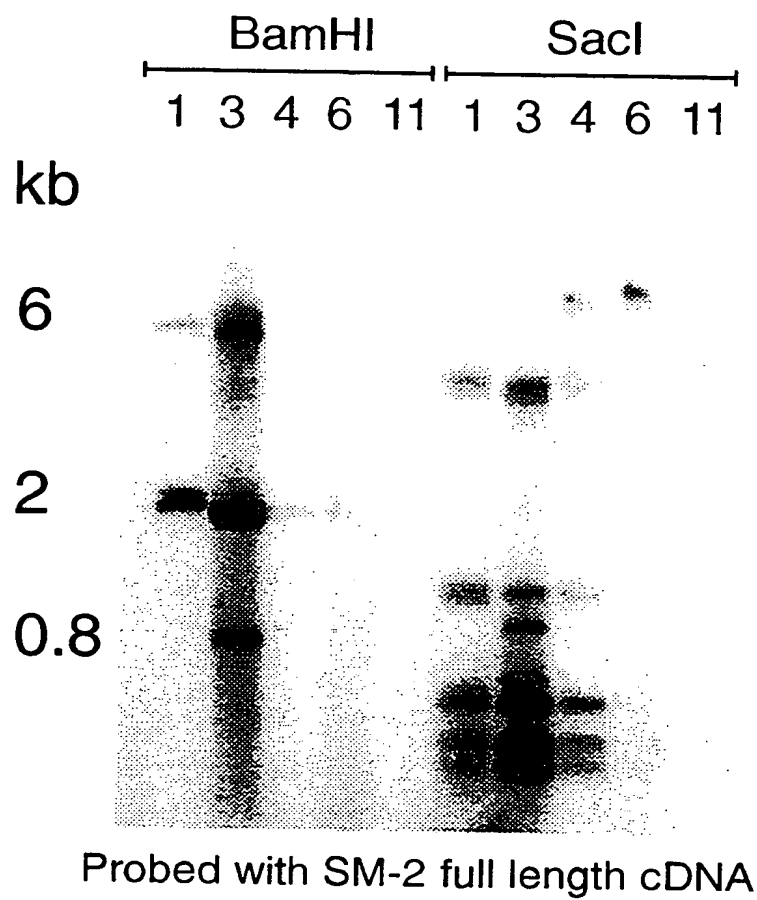


FIGURE 17

28/44

## INTRON EXON STRUCTURE - Wheat SSI

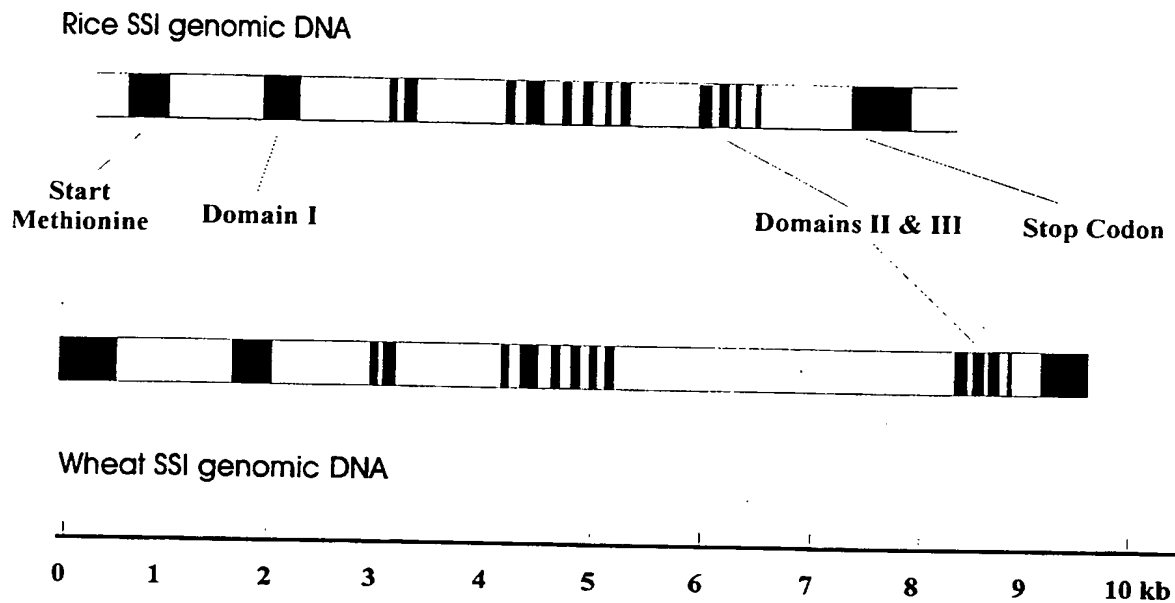
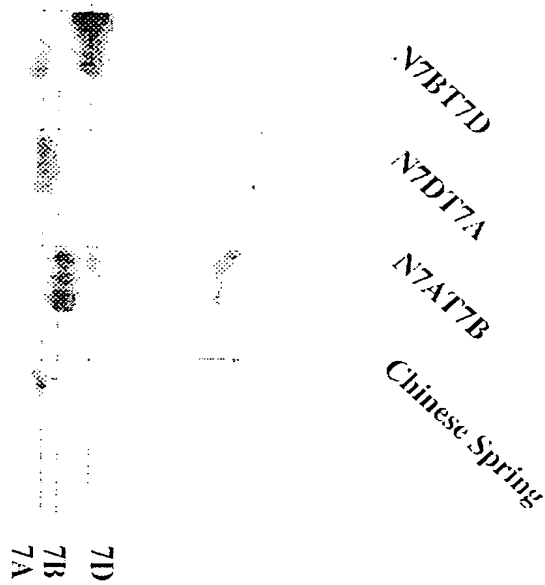


FIGURE 18

29 / 44



**FIGURE 19**

30/44

```

      ATACTACATACTATATGCTTGCACCCAGGACACTTTTATAACTATTCTGGCTGTGGGA
80  +-----+-----+-----+-----+-----+-----+-----+-----+
      TATGATGTATGATATACGAACGTGGTCCCTGTGAAAATATTGATAAGACCGACACCCCT
      139
a      T T Y Y M L A P K G H F Y N Y S G C G N -
b      I L H T I C L H P R D T F I T I L A V G -
c      Y Y I L Y A C T Q G T L L * L F W L W E -

      ATACCTTCAACTGTAATCATCCTGTGGTTCGTCATTCATTGTAGATTGTTAAGATACT
140 +-----+-----+-----+-----+-----+-----+-----+-----+
      TATGGAAGTTGACATTAGTAGGACACCAAGCAGTTAAGTAACATCTAACAATTCATATGA
      199
a      T F N C N H P V V R Q F I V D C L R Y W -
b      I P S T V I I L W F V N S L * I V * D T -
c      Y L Q L * S S C G S S I H C R L F K I L -

      GGGTGACGGAAATGCATGTTGATGGTTTCGTTTGTGACCTT
200 +-----+-----+-----+-----+-----+-----+-----+-----+
      CCCACTGCCCTTTACGTACAACACTACCAAAAGCAAACTGGAA
      240
a      V T E M H V D G F R F D L -
b      G * R K C M L M V F V L T -
c      G D G N A C * W F S F * P -

Enzymes that do cut:

      NONE

Enzymes that do not cut:

      EcoRI
  
```

Figure 20a

31/44

Comparison of Wheat Debranching Enzyme-I (WDBE-I) PCR fragment with maize *Sugary-1* DNA sequence

SUGARY.DNA	1098	1107	1117	1127	1137	1147	1157
		TGAGGTGATCATGGATGTTGTCCTTCATCATACAGCTGAGGTAATGAGAAAGGCCCAAT					
WHEAT1.DNA							
		....GTGATCATGGATGTTGTCCTTCACACCATACAGCTGAGGTAATGAGAAATGGTCCAAT					
-3	6	16	26	36	46	56	
FILE NAME	1158	1167	1177	1187	1197	1207	1217
SUGARY.DNA		ATTATCCTTTAGGGGATAGATAAATAGTACATACACTGCTTGACCTTAAGGAGAGTT					
WHEAT1.DNA							
		ATTATCATTTAGGGGTCGATAAATACATACATATATGCTTGACCCCAAGGACACTT					
57	66	76	86	96	106	116	
FILE NAME	1218	1227	1237	1247	1257	1267	1277
SUGARY.DNA		TTATAATTATCTGGTTGTGGAATACCTTCAATTGTAATCATCTGTTAGTCCGTGAATT					
WHEAT1.DNA							
		TTATAACTATTCTGGCTGTGGGNATACCTTCAACTGTAATCATCTGTTGGTTCGTCAATT					
117	126	136	146	156	166	176	
FILE NAME	1278	1287	1297	1307	1317	1327	1337
SUGARY.DNA		TATAGTGGATTGCTTGAGATACTGGGTACAGAAATGCATGTTGATGGTTTCGTTTGA					
WHEAT1.DNA							
		CATTGTAGATTGTTAAGNTACTGGGTGACGGAAATGCATGTTGTTTCGTTTGA					
177	186	196	206	216	226	236	
FILE NAME	1338	1347	1357				
SUGARY.DNA		CCTTGCACTATACT-G...					
WHEAT1.DNA							
		CCTTGCACTN--CTNAAA					
237	246	256					
MATCHING PERCENTAGE							
TOTAL WINDOW		84%	( 219/ 260)				
ALIGNMENT WINDOW		86%	( 219/ 253)				

Figure 20b

32/44

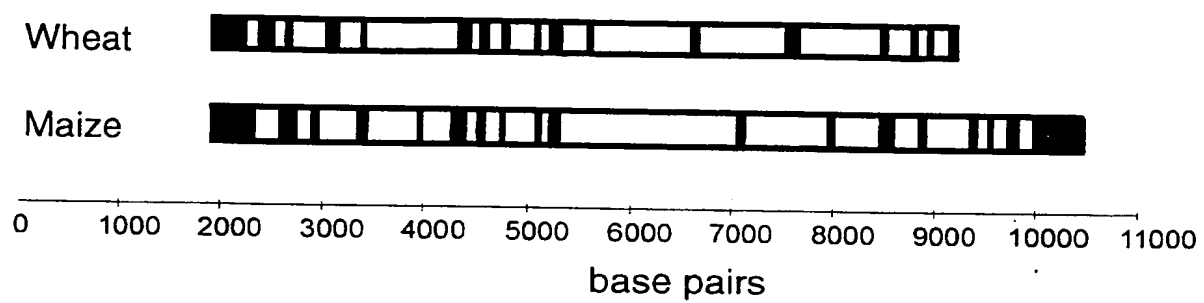


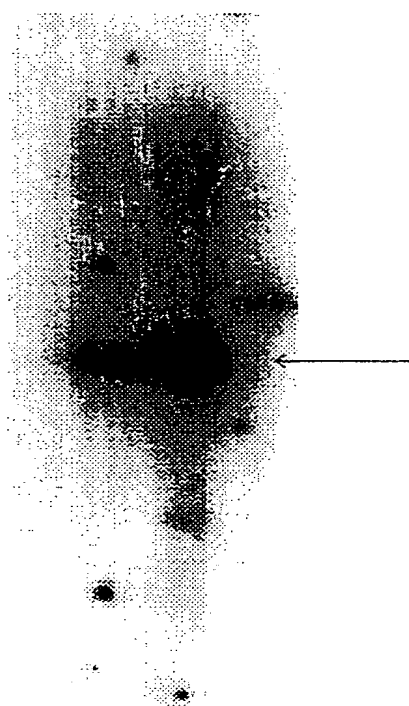
FIGURE 20C



33 / 44

Southern blot of *T. tauschii*  
Genomic DNA

1X 3X



BamHI Digest

*T. tauschii* Genomic DNA Probed  
With The Wheat Debranching Enzyme  
PCR Product

FIGURE 21A

34/44

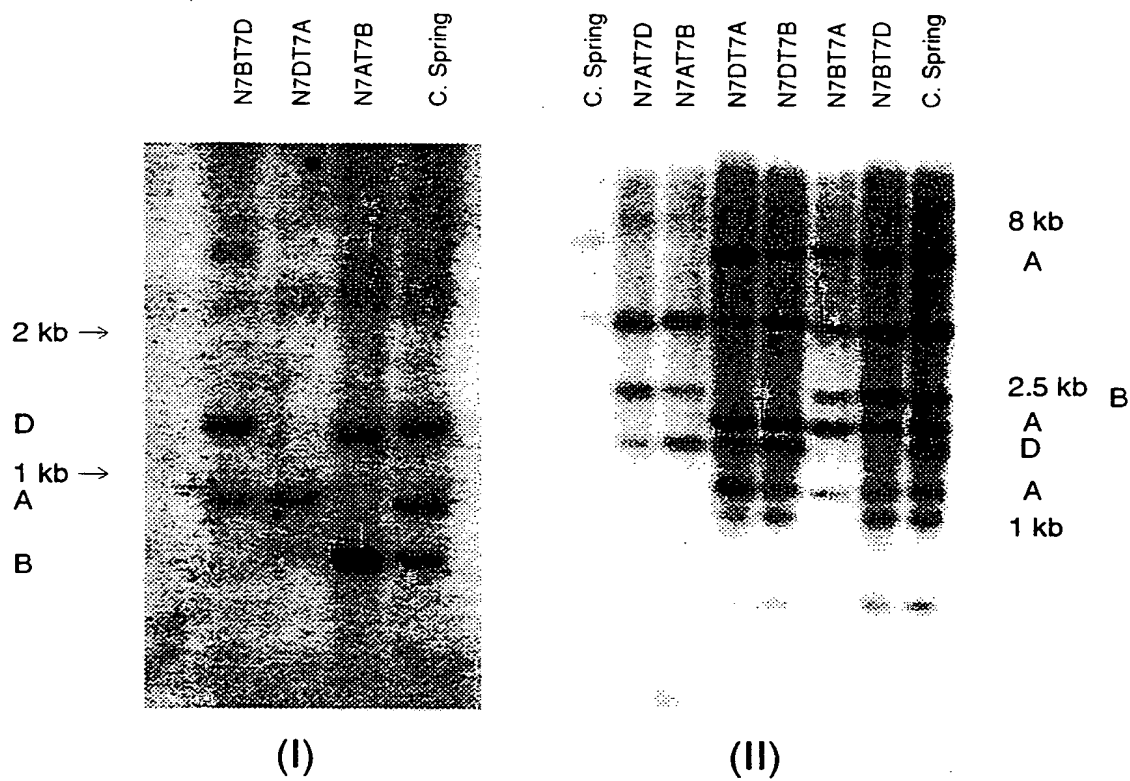


FIGURE 21B

35 / 44

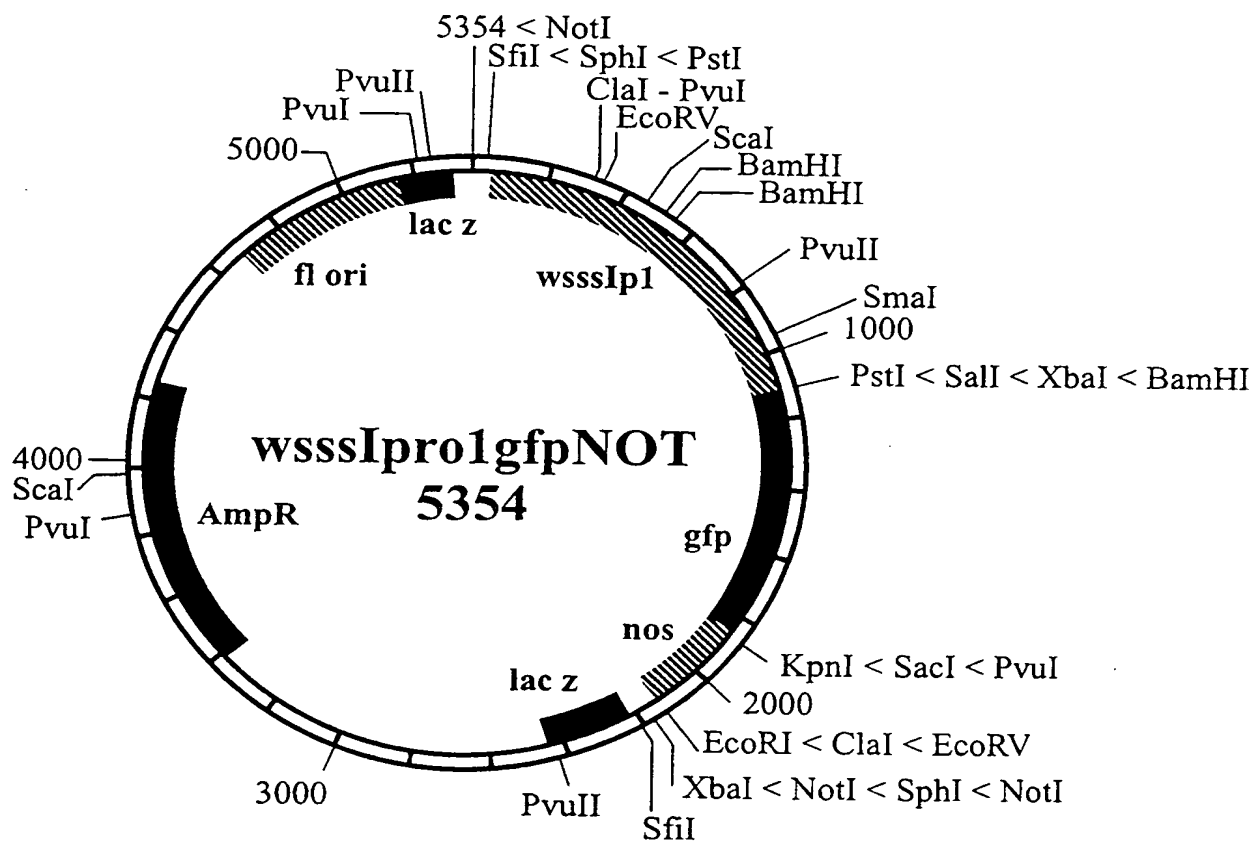


FIGURE 22A

36 / 44

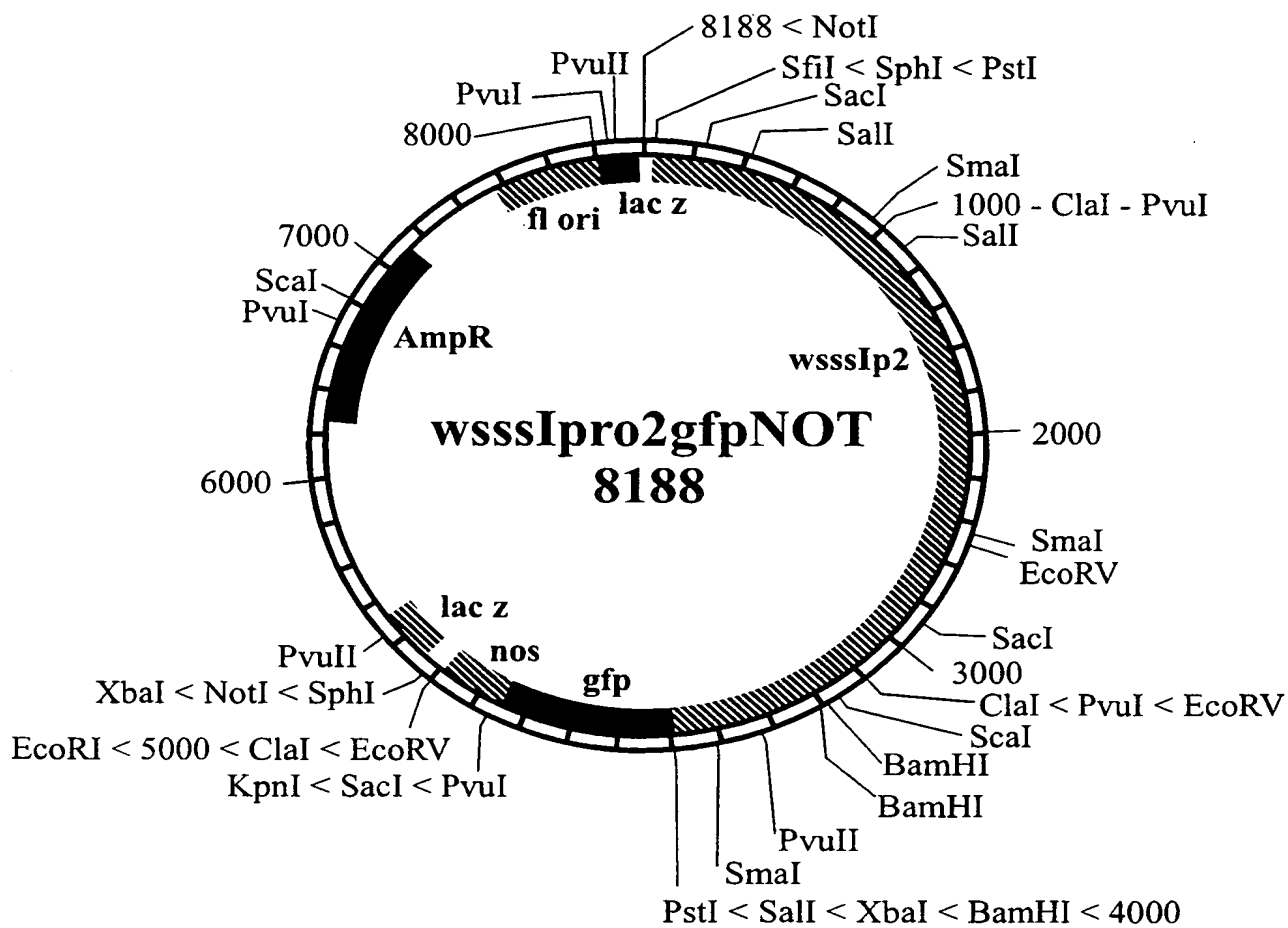


FIGURE 22B

37 / 44

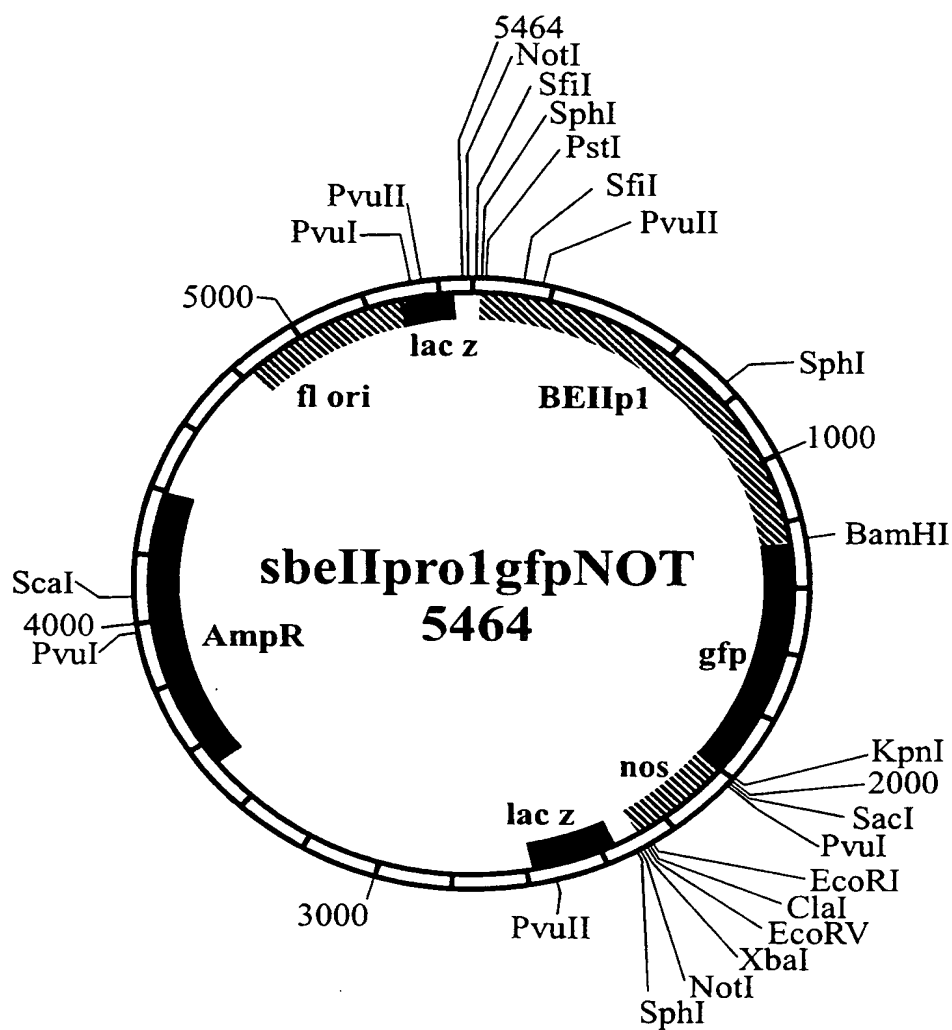


FIGURE 22C

38 / 44

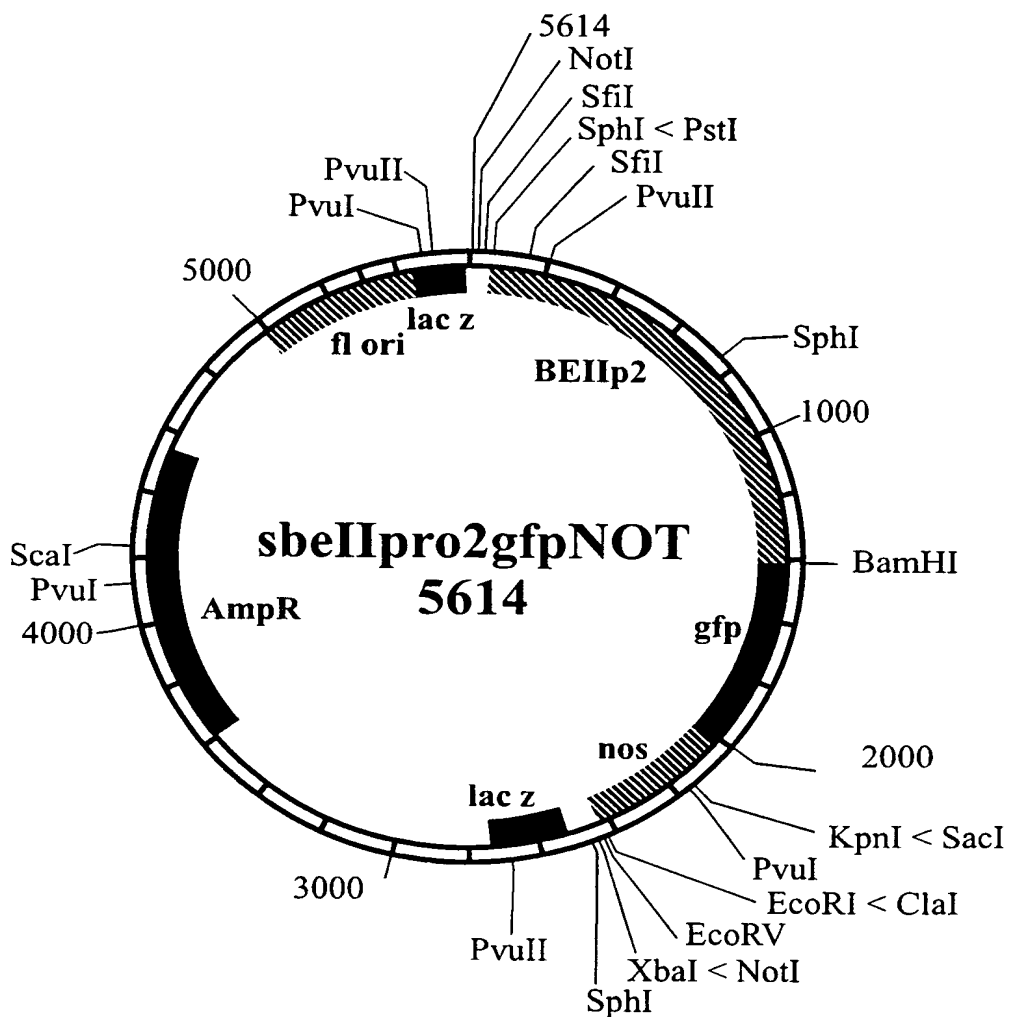
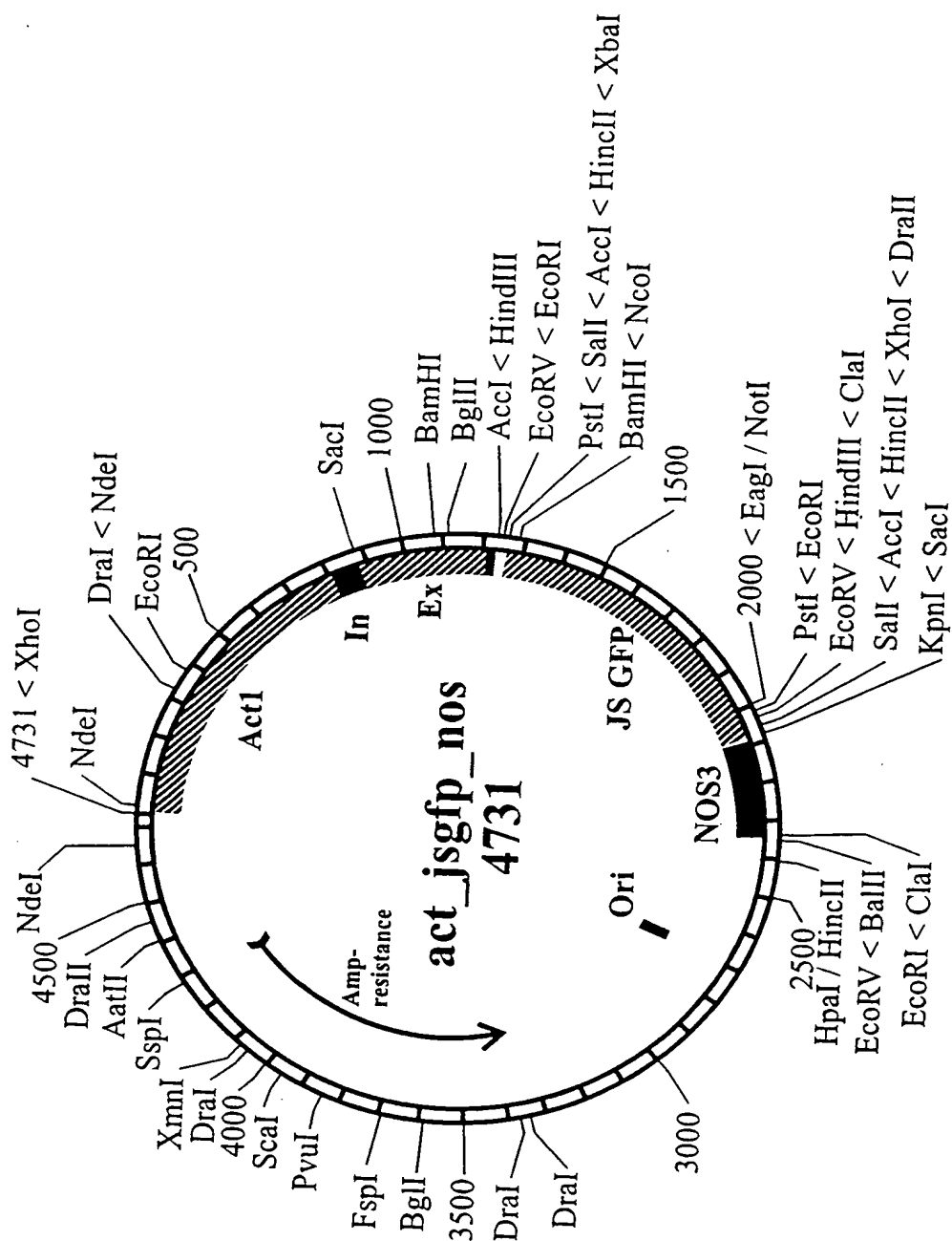
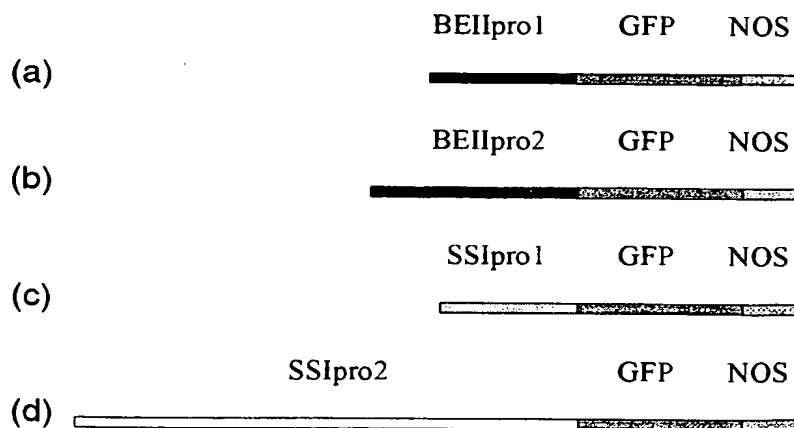
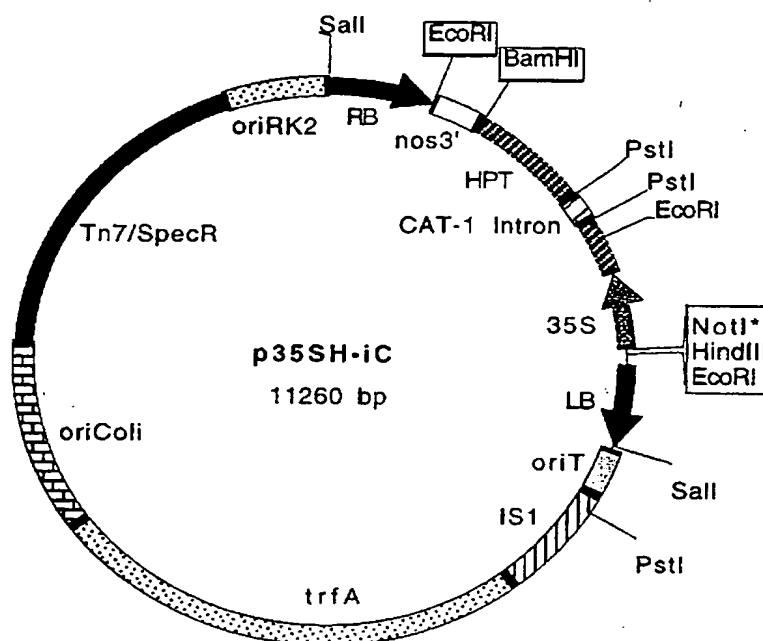


FIGURE 22D

39 / 44



**Figure 22E**  
 SUBSTITUTE SHEET (Rule 26) (RO/AU)



**FIGURE 23**



41/44

Primer Set	Key	Forward Primer	Forward Primer Sequence
1	E01'/E02	WBE2E1F	CGT CGC TGC TCC TCA GGA AG
2	E01/E02	sr854.1180F	CTG GCT GAC TCA ATC ACT ACG
3	E02/E03	WBE2E2F	CGC AAC CTG AAG AAT TAC AG
4	E03/E04	WBE2E3F	ATT TTC GGA GCC ATC TTG AC
5	E04/E05	WBE2E4F	TCG TGG TTA TGA AAA GCT TGG
6	E05/E06	sr913F	ATC ACT TAC CGA GAA TGG G
7	E05/I05	sr913F	ATC ACT TAC CGA GAA TGG G
8	E06/E07	WBE2E6F	ACA ATT GGA ATC CAA ATG CA
9	E07/E08	WBE2E7F	AGC TAT TCC TCA TGG CTC AC
10	E08/E09	WBE2E8F	TGC AGG CTC CAG GTG AAA TA
11	E10/E11	da5.seq	GGC TTG GAT ACA ATG CAG TGC
12	E12/E13	da151.seq	TTG ACG GCT TGA ATG GTT TC
13	E17/E18	WBE2E17F	TTT AGG TGG TGA AGG CTA TCT
14	E18/E19	sr860R	AAT GGA TAG ATT TTC CAA GAG G
15	E19_3'	WBE2-2395F	AGC AGA ACT GCG GTC GTG TA

Reverse Primer	Reverse Primer Sequence	Temp	bp
WBE2E2R	CAG GAC CTT CCC TGG AGA GG	57.4	401
WSBE9E2R	GGC ACG AGT GTG TGT ACC TGT A	57.7	601
sr866F	TAT CTT CAG GTA TCT ACA GC	49.8	309
WBE2E4R2	ATG CTT CCA ATC CAC CTT CA	-	>450
WBE2E5R	GAG CCC ATT CTC GGT AAG TGA	50.5	234
WBE2E6R	CTG CAT TTG GAT TCC AAT TG	49.9	232
WBE2I5R	CAG TAA GCT AGT TGG TGA ATA	46.6	106
WBE2E7R	GGG AGG AAA ATC TCC CAA AC	51.0	402
sr915F	CCA TTG AAA GGT ATT TCA CC	51.1	203
sr912F	TAA CTT ATT GAC ATA CCG G	48.4	439
WBE2E11R	CTG GAG TTC CAA AAC GGC TAC	51.2	289
WBE2E13R	ATT CTT CAA GCC ACC ATC TC	51.6	244
WBE2E18R	TAT TGT TAT TTC CAG GGG AGA	50.2	258
da23.seq	TGC TGC ATT GCC TGA TCG AA	50.4	~295
WBE2-2634R	AAC ACC CAG GCC CGT CCA TT	57.2	240

Figure 24

42 / 44

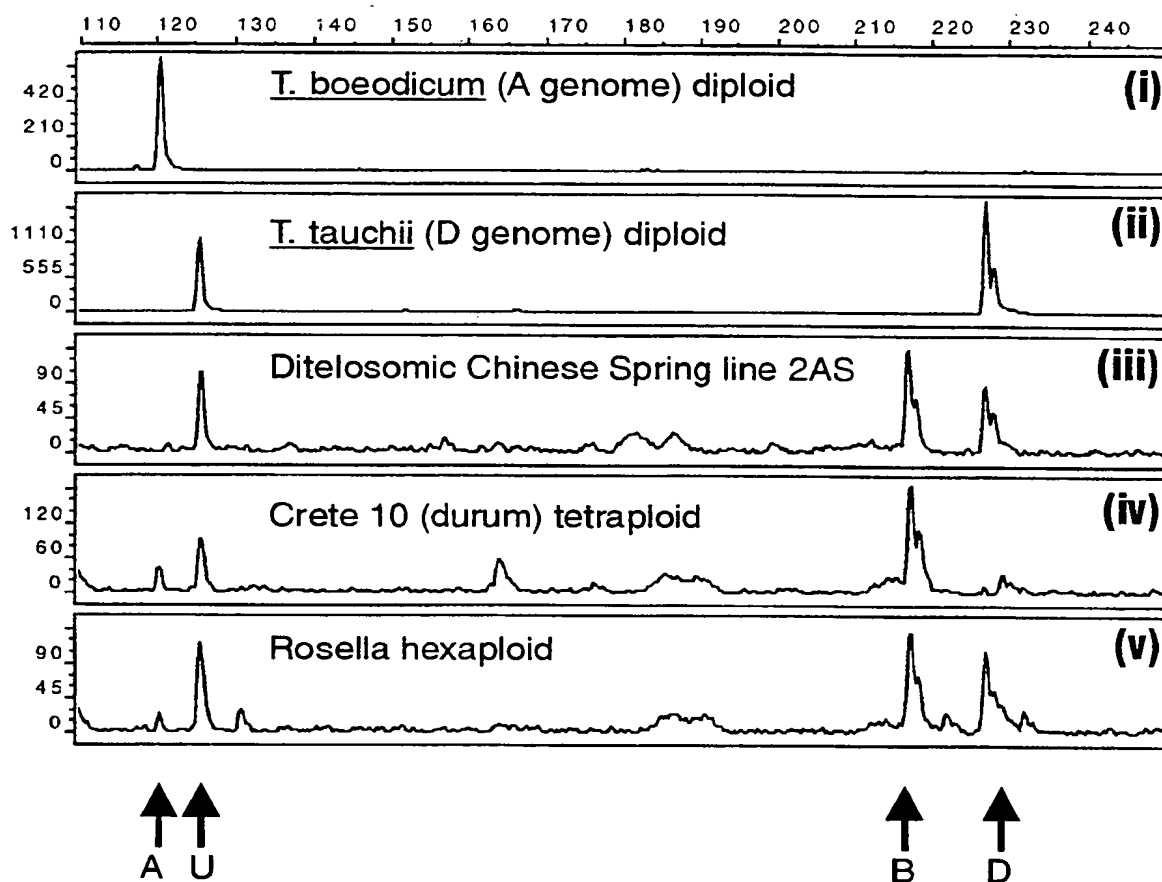
SBE II Intron 5 primer set - digested with DdeI

FIGURE 25

43 / 44

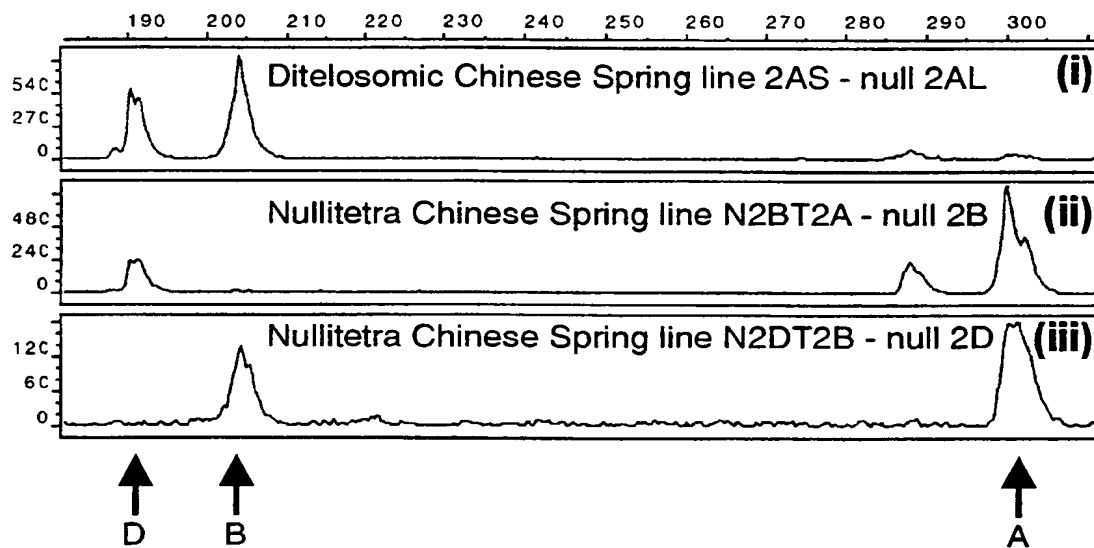
SBE II Intron 10 primer set - digested with DdeI

FIGURE 26

44 / 44

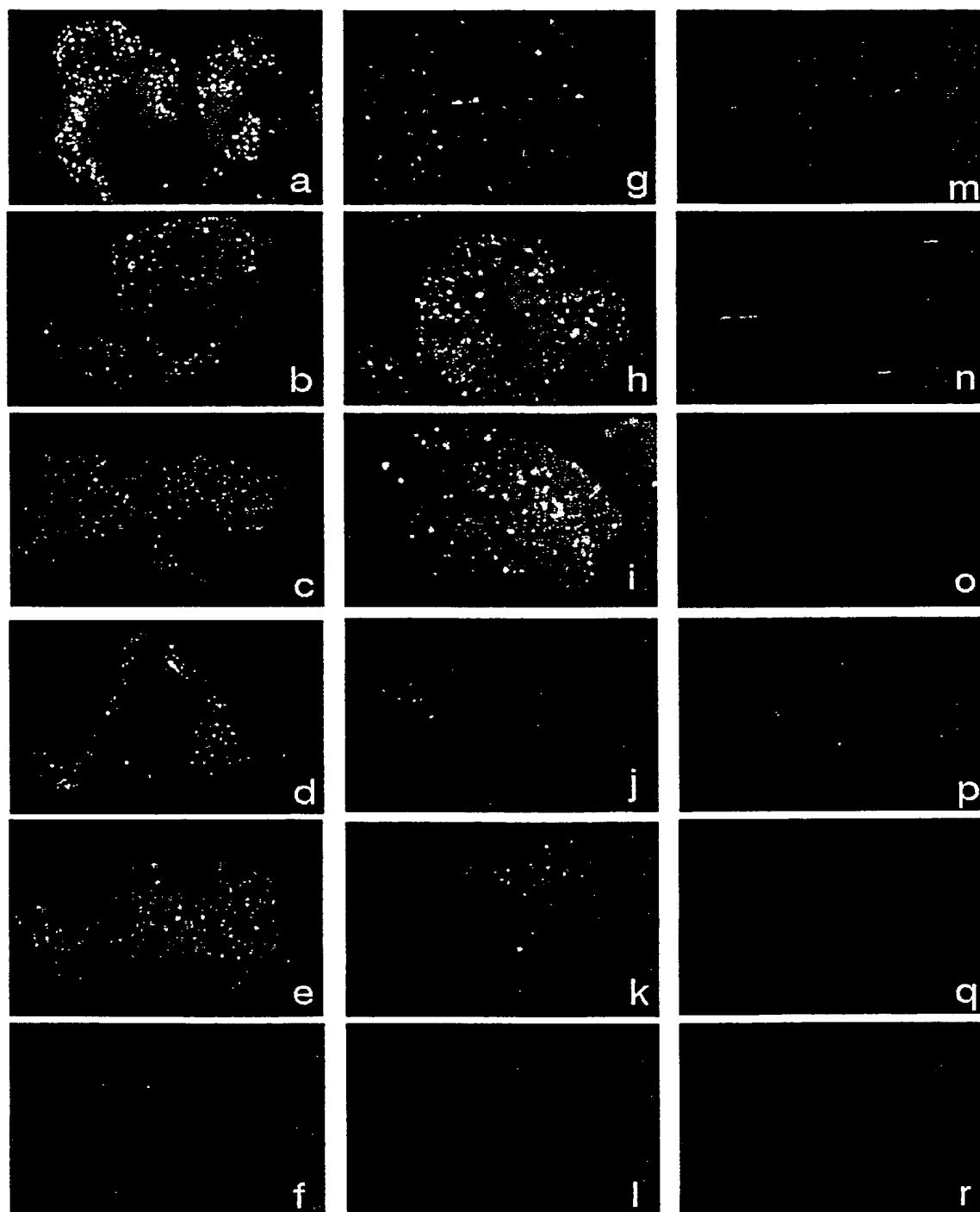


FIGURE 27

CORRECTED VERSION

PATENT COOPERATION TREATY

PCT

INTERNATIONAL SEARCH REPORT

(PCT Article 18 and Rules 43 and 44)

Applicant's or agent's file reference	<b>FOR FURTHER ACTION</b> see Notification of Transmittal of International Search Report (Form PCT/ISA/220) as well as, where applicable, item 5 below.	
International application No. <b>PCT/AU 98/00743</b>	International filing date ( <i>day/month/year</i> ) <b>11 September 1998</b>	(Earliest) Priority Date ( <i>day/month/year</i> ) <b>12 September 1997</b>
Applicant (1) Commonwealth Scientific and Industrial Research Organisation et al (2) LI, Zhongyi et al		

This international search report has been prepared by this International Searching Authority and is transmitted to the applicant according to Article 18. A copy is being transmitted to the International Bureau.

This international search report consists of a total of **4** sheets.

☒ It is also accompanied by a copy of each prior art document cited in this report.

**1. Basis of the report**

- a. With regard to the language, the international search was carried out on the basis of the international application in the language in which it was filed, unless otherwise indicated under this item.

☐ the international search was carried out on the basis of a translation of the international application furnished to this Authority (Rule 23.1(b)).

- b. With regard to any nucleotide and/or amino acid sequence disclosed in the international application, the international application, the international search was carried out on the basis of the sequence listing:

☒ contained in the international application in written form.

☐ filed together with the international application in computer readable form.

☐ furnished subsequently to this Authority in written form.

☐ furnished subsequently to this Authority in computer readable form.

☐ the statement that the subsequently furnished written sequence listing does not go beyond the disclosure in the international application as filed has been furnished.

☐ the statement that the information recorded in computer readable form is identical to the written sequence listing has been furnished

2. ☐ Certain claims were found unsearchable (See Box I).

3. ☐ Unity of invention is lacking (See Box II).

4. With regard to the title, ☒ the text is approved as submitted by the applicant.

☐ the text has been established by this Authority to read as follows:

5. With regard to the abstract, ☒ the text is approved as submitted by the applicant

☐ the text has been established, according to Rule 38.2(b), by this Authority as it appears in Box III. The applicant may, within one month from the date of mailing of this international search report, submit comments to this Authority.

6. The figure of the drawings to be published with the abstract is Figure No.

☐ as suggested by the applicant.

☐ because the applicant failed to suggest a figure

☐ because this figure better characterizes the invention

☒ None of the figures

# INTERNATIONAL SEARCH REPORT

International application No.

PCT/AU 98/00743

## A. CLASSIFICATION OF SUBJECT MATTER

Int Cl<sup>6</sup>: C12N 9/24, 15/55

According to International Patent Classification (IPC) or to both national classification and IPC

## B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

See Electronic Data base box

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

See Electronic Data base box

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

WPAT - Starch branching enzyme #, promoter #, debranching enzyme : CA, medline - Starch Branching enzyme #, starch synthase, triticum, wheat: Genebank, Embl - sequences as claimed.

## C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	AU-B-19028/95 (688006) (Nat. Starch & Chem) 17 October 1995. (See fig 8 in particular)	1, 2, 16, 21, 22 and 36
PX	AU-A 48747/97 (Nat. Starch & Chem) 14 May 1998. Epd 5 November 1996 (See Fig 4 in particular)	1, 2, 16, 21, & 22
X	WO 97/04113 (DANISCO A/S) 6 February 1997 (See fig 8 and page 22 in particular)	1, 2, 16, 21& 22

☒ Further documents are listed in the continuation of Box C

☒ See patent family annex

### \* Special categories of cited documents:

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier application or patent but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

"&" document member of the same patent family

Date of the actual completion of the international search

13 October 1998

Date of mailing of the international search report

20 OCT 1998

Name and mailing address of the ISA/AU

AUSTRALIAN PATENT OFFICE

PO BOX 200

WODEN ACT 2606

AUSTRALIA

Facsimile No.: (02) 6285 3929

Authorized officer

PHILIPPA WYRDEMAN

Telephone No.: (02) 6283 2554

# INTERNATIONAL SEARCH REPORT

International application No.

PCT/AU 98/00743

C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT		
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	AU-B-65392/94 (693787) (DANISCO A/S) 8 November 1994. (See page 43 in particular)	1, 2, 16, 21 & 22
X	AU - A 77165/95 (AMYLOGENE HB) 5 June 1997 (See in particular seq. ID# 1, page 12)	1, 2, 16, 21 & 22
X	Nair, R. B et al (1997) <u>PLANT SCIENCE</u> "Isolation, characterisation and expression analysis of a starch branching enzyme II cDNA from wheat" vol. 122, pages 153-163. (See entire document)	1, 2, 16, 21, & 22

# INTERNATIONAL SEARCH REPORT

## Information on patent family members

International application No.  
PCT/AU 98/00743

This Annex lists the known "A" publication level patent family members relating to the patent documents cited in the above-mentioned international search report. The Australian Patent Office is in no way liable for these particulars which are merely given for the purpose of information.

Patent Document Cited in Search Report				Patent Family Member			
WO	9704113	AU	66146/96	EP	839203		
AU	94/65392	CA	2160159	EP	693128	GB	2291878
		NZ	265061	WO	9424292		
AU	95/77165	WO	97/20040	EP	863983	NO	982443
		SE	9601506	SE	9504272		
AU	95/19028	WO	9526407	EP	754235	CA	2186399
AU	97/48747	WO	9820145	GB	2320716		
GB	9307408	AU	65392/94	CA	2160159	EP	693128
		GB	2291878	NZ	265061	WO	9424292
SE	9504272	AU	77165/96	EP	863983	NO	982443
		SE	9601506	WO	9720040		
GB	9406022	AU	19028/95	CA	2186399	EP	754235
		WO	9526407				
GB	9623095	AU	48747/97	GB	2320716	WO	9820145
END OF ANNEX							